

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 14:23:32 ; Search time 82 Seconds  
(without alignments)  
135.354 Million cell updates/sec

Title: US-09-825-489-3  
Perfect score: 20  
Sequence: 1 ggtccatccatcgttgatg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 800 summaries

Database : Issued\_Patents\_NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	100.0	822	4 US-09-651-656-8	Sequence 8, Appl1
C 2	20	100.0	822	4 US-09-650-855-8	Sequence 8, Appl1
C 3	16.8	84.0	14273	4 US-08-961-527-40	Sequence 40, Appl1
C 4	15.8	79.0	614	3 US-09-385-988-62	Sequence 62, Appl1
C 5	15.8	79.0	686	4 US-09-023-655-1225	Sequence 1225, Ap
C 6	15.8	79.0	801	4 US-09-543-681A-137	Sequence 137, Ap
C 7	15.8	79.0	1176	4 US-09-543-681A-1773	Sequence 1773, Ap
C 8	15.4	77.0	675	4 US-09-134-001C-388	Sequence 388, Ap
C 9	15.2	76.0	330	4 US-08-956-171E-1137	Sequence 1137, Ap
C 10	15.2	76.0	1419	4 US-09-489-039A-5224	Sequence 5224, Ap
C 11	15.2	76.0	4884	4 US-09-328-352-2478	Sequence 2478, Ap
C 12	15.2	76.0	14311	3 US-08-646-695-1	Sequence 1, Appl1
C 13	15.2	76.0	14311	3 US-08-646-695-7	Sequence 7, Appl1
C 14	15.2	76.0	14311	5 PCT-US96-06053-1	Sequence 1, Appl1
C 15	15.2	76.0	14311	5 PCT-US96-06053-7	Sequence 7, Appl1
C 16	15.2	76.0	36412	4 US-08-311-731A-132	Sequence 132, App
C 17	15.2	76.0	1230025	4 US-09-198-452A-1	Sequence 1, Appl1
C 18	15.2	76.0	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
C 19	15.2	76.0	4411529	3 US-09-103-840A-1	Sequence 1, Appl1
C 20	14.8	74.0	315	4 US-09-328-352-1128	Sequence 1128, Ap
C 21	14.8	74.0	348	4 US-09-328-352-1227	Sequence 1227, Ap
C 22	14.8	74.0	387	4 US-09-621-976-3741	Sequence 3741, Ap
C 23	14.8	74.0	546	4 US-09-328-352-992	Sequence 992, App
C 24	14.4	72.0	2244	4 US-09-489-039A-911	Sequence 911, App
C 25	14.4	72.0	2310	4 US-09-489-039A-953	Sequence 953, App
C 26	14.4	72.0	2433	1 US-08-136-743B-40	Sequence 40, Appl
C 27	14.4	72.0	2663	1 US-08-136-743B-3	Sequence 3, Appl1

28	14.4	72.0	50341	1 US-08-247-901C-1	Sequence 1, Appl1
29	14.4	72.0	50341	2 US-09-075-904-1	Sequence 1, Appl1
30	14.4	72.0	52297	4 US-09-426-436-1	Sequence 1, Appl1
31	14.4	72.0	52297	4 US-08-705-557-1	Sequence 1, Appl1
32	14.2	71.0	36	3 US-09-135-782-10	Sequence 10, Appl
33	14.2	71.0	329	4 US-08-956-171E-986	Sequence 986, App
C 34	14.2	71.0	378	4 US-09-134-001C-2448	Sequence 2448, Ap
C 35	14.2	71.0	423	4 US-09-252-991A-11719	Sequence 11719, A
C 36	14.2	71.0	480	4 US-09-252-991A-13151	Sequence 13151, A
C 37	14.2	71.0	510	4 US-09-134-000C-244	Sequence 244, App
C 38	14.2	71.0	555	4 US-09-107-532A-735	Sequence 735, App
C 39	14.2	71.0	743	4 US-09-342-681C-11	Sequence 11, Appl
C 40	14.2	71.0	909	4 US-09-252-991A-9153	Sequence 9153, Ap
C 41	14.2	71.0	1065	4 US-09-634-238-121	Sequence 121, App
C 42	14.2	71.0	1176	4 US-09-342-681C-14	Sequence 14, Appl
C 43	14.2	71.0	1185	4 US-09-252-991A-11577	Sequence 11577, A
C 44	14.2	71.0	1225	1 US-08-286-020-1	Sequence 1, Appl1
C 45	14.2	71.0	1225	1 US-08-603-919-1	Sequence 1, Appl1
C 46	14.2	71.0	1315	2 US-08-343-101A-8	Sequence 8, Appl1
C 47	14.2	71.0	1315	3 US-09-183-688-8	Sequence 8, Appl1
C 48	14.2	71.0	1315	4 US-09-519-689-8	Sequence 8, Appl1
C 49	14.2	71.0	1386	4 US-09-252-991A-11863	Sequence 11863, A
C 50	14.2	71.0	1398	4 US-09-489-039A-3739	Sequence 3739, Ap
C 51	14.2	71.0	1419	4 US-09-252-991A-11636	Sequence 11636, A
C 52	14.2	71.0	1574	4 US-09-342-681C-1	Sequence 1, Appl1
C 53	14.2	71.0	1651	4 US-09-533-029-59	Sequence 59, Appl
C 54	14.2	71.0	1654	4 US-09-634-238-16	Sequence 16, Appl
C 55	14.2	71.0	1707	4 US-09-134-001C-931	Sequence 931, App
C 56	14.2	71.0	1845	4 US-09-307-973A-1	Sequence 1, Appl1
C 57	14.2	71.0	1875	4 US-09-252-991A-8812	Sequence 8812, Ap
C 58	14.2	71.0	1901	5 PCT-US93-05000-32	Sequence 32, Appl
C 59	14.2	71.0	1954	4 US-09-023-655-286	Sequence 286, App
C 60	14.2	71.0	1989	4 US-09-377-466B-7	Sequence 7, Appl1
C 61	14.2	71.0	1989	4 US-09-377-466B-7	Sequence 7, Appl1
C 62	14.2	71.0	2000	3 US-08-996-441B-99	Sequence 99, Appl
C 63	14.2	71.0	2000	3 US-08-993-722A-99	Sequence 99, Appl
C 64	14.2	71.0	2000	3 US-08-993-722A-99	Sequence 99, Appl
C 65	14.2	71.0	2000	3 US-08-993-722A-99	Sequence 99, Appl
C 66	14.2	71.0	2000	4 US-09-427-770-99	Sequence 99, Appl
C 67	14.2	71.0	2000	4 US-09-427-770-99	Sequence 99, Appl
C 68	14.2	71.0	2022	2 US-08-464-517-32	Sequence 32, Appl
C 69	14.2	71.0	2022	2 US-08-246-361A-32	Sequence 32, Appl
C 70	14.2	71.0	2022	2 US-08-463-772-32	Sequence 32, Appl
C 71	14.2	71.0	2050	3 US-08-996-441B-101	Sequence 101, App
C 72	14.2	71.0	2050	3 US-08-993-722A-101	Sequence 101, App
C 73	14.2	71.0	2050	3 US-08-993-722A-101	Sequence 101, App
C 74	14.2	71.0	2050	3 US-08-993-722A-101	Sequence 101, App
C 75	14.2	71.0	2050	4 US-09-427-770-101	Sequence 101, App
C 76	14.2	71.0	2050	4 US-09-427-770-101	Sequence 101, App
C 77	14.2	71.0	2152	1 US-08-188-582-17	Sequence 17, Appl
C 78	14.2	71.0	2152	1 US-08-646-715-17	Sequence 17, Appl
C 79	14.2	71.0	2211	3 US-09-318-448-26	Sequence 26, Appl
C 80	14.2	71.0	2211	4 US-09-347-878-2	Sequence 2, Appl1
C 81	14.2	71.0	2211	4 US-09-546-013-4	Sequence 4, Appl1
C 82	14.2	71.0	2259	1 US-08-420-235B-20	Sequence 20, Appl
C 83	14.2	71.0	2259	1 US-08-793-654-20	Sequence 20, Appl
C 84	14.2	71.0	2259	5 PCT-US95-10194-20	Sequence 20, Appl
C 85	14.2	71.0	2304	4 US-09-252-991A-11802	Sequence 11802, A
C 86	14.2	71.0	2344	4 US-07-695-472B-35	Sequence 35, Appl
C 87	14.2	71.0	2344	4 US-09-106-375-35	Sequence 35, Appl
C 88	14.2	71.0	2347	4 US-08-250-740-32	Sequence 32, Appl
C 89	14.2	71.0	2347	1 US-07-695-472B-34	Sequence 1, Appl1
C 90	14.2	71.0	2347	1 US-07-695-472B-34	Sequence 34, Appl
C 91	14.2	71.0	2347	1 US-07-695-472B-36	Sequence 36, Appl
C 92	14.2	71.0	2347	4 US-09-106-375-1	Sequence 1, Appl1
C 93	14.2	71.0	2347	4 US-09-106-375-34	Sequence 34, Appl1
C 94	14.2	71.0	2347	4 US-09-106-375-36	Sequence 36, Appl
C 95	14.2	71.0	2373	4 US-09-023-655-1011	Sequence 1011, Ap
C 96	14.2	71.0	2451	1 US-07-731-137A-3	Sequence 3, Appl1
C 97	14.2	71.0	2451	1 US-08-229-444B-1	Sequence 1, Appl1
C 98	14.2	71.0	2451	2 US-08-541-780-3	Sequence 3, Appl1
C 99	14.2	71.0	2517	4 US-09-252-991A-9238	Sequence 9238, Ap
C 100	14.2	71.0	2549	1 US-08-470-720-2	Sequence 2, Appl1

101	14.2	71.0	2706	4	US-09-134-001C-2552	Sequence 2552, Ap	C 174	13.8	69.0	3797	1	US-08-789-449-1	Sequence 1, Appli
C 102	14.2	71.0	3043	4	US-09-049-698-16	Sequence 16, Appl	C 175	13.8	69.0	11707	3	US-09-136-574A-1	Sequence 1, Appli
C 103	14.2	71.0	3044	4	US-09-377-466B-38	Sequence 38, Appl	C 176	13.8	69.0	28626	4	US-09-596-002-15	Sequence 15, Appli
C 104	14.2	71.0	3181	4	US-09-049-698-18	Sequence 18, Appl	C 177	13.8	69.0	36682	3	US-08-943-731-2	Sequence 2, Appli
C 105	14.2	71.0	3198	4	US-09-345-236B-83	Sequence 83, Appl	C 178	13.8	69.0	43676	3	US-09-356-952-12	Sequence 12, Appli
C 106	14.2	71.0	3198	4	US-09-345-236B-87	Sequence 87, Appl	C 179	13.8	69.0	43312	4	US-09-671-317-485	Sequence 485, App
C 107	14.2	71.0	3198	4	US-09-345-236B-90	Sequence 90, Appl	C 180	13.8	69.0	63588	4	US-09-873-404-3	Sequence 3, Appli
C 108	14.2	71.0	3198	4	US-09-345-236B-92	Sequence 92, Appl	C 181	13.8	69.0	1230025	4	US-09-198-452A-1	Sequence 1, Appli
C 109	14.2	71.0	3362	2	US-08-678-039A-41	Sequence 41, Appl	C 182	13.8	69.0	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 110	14.2	71.0	3455	4	US-09-377-466B-36	Sequence 36, Appl	C 183	13.8	69.0	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 111	14.2	71.0	3518	4	US-09-412-210-2	Sequence 2, Appli	C 184	13.6	68.0	115	4	US-09-494-921-20	Sequence 20, Appl
C 112	14.2	71.0	3518	4	US-10-121-911A-2	Sequence 2, Appli	C 185	13.6	68.0	272	4	US-09-016-434-95	Sequence 439, App
C 113	14.2	71.0	3500	4	US-09-134-000C-2137	Sequence 2137, Ap	C 186	13.6	68.0	289	4	US-09-313-294A-2910	Sequence 2910, Ap
C 114	14.2	71.0	3754	4	US-09-377-466B-15	Sequence 15, Appl	C 187	13.6	68.0	315	4	US-09-107-532A-2050	Sequence 2050, Ap
C 115	14.2	71.0	4008	4	US-09-134-001C-879	Sequence 879, App	C 188	13.6	68.0	336	4	US-09-328-352-3353	Sequence 3353, Ap
C 116	14.2	71.0	4149	4	US-09-377-466B-13	Sequence 13, Appl	C 189	13.6	68.0	400	4	US-08-956-111E-3987	Sequence 3897, Ap
C 117	14.2	71.0	4223	4	US-09-620-312D-857	Sequence 857, App	C 190	13.6	68.0	406	4	US-08-956-111E-3944	Sequence 3944, Ap
C 118	14.2	71.0	4381	4	US-09-347-878-19	Sequence 19, Appl	C 191	13.6	68.0	409	4	US-09-621-976-3944	Sequence 3157, Ap
C 119	14.2	71.0	4741	1	US-07-695-472B-4	Sequence 4, Appli	C 192	13.6	68.0	410	4	US-09-494-921-6	Sequence 6, Appli
C 120	14.2	71.0	4741	1	US-09-106-375-4	Sequence 4, Appli	C 193	13.6	68.0	430	4	US-09-621-976-3161	Sequence 7, Appli
C 121	14.2	71.0	4742	1	US-08-250-740-35	Sequence 35, Appli	C 194	13.6	68.0	451	4	US-09-621-976-3161	Sequence 3161, Ap
C 122	14.2	71.0	4748	3	US-09-331-581-1	Sequence 1, Appli	C 195	13.6	68.0	456	4	US-09-107-532A-2457	Sequence 2457, Ap
C 123	14.2	71.0	4851	4	US-09-221-017B-999	Sequence 999, App	C 196	13.6	68.0	473	4	US-09-621-976-856	Sequence 856, App
C 124	14.2	71.0	4964	1	US-08-470-720-5	Sequence 5, Appli	C 197	13.6	68.0	494	4	US-09-621-976-3160	Sequence 3160, Ap
C 125	14.2	71.0	6157	4	US-08-956-171E-184	Sequence 184, App	C 198	13.6	68.0	528	4	US-09-621-976-3159	Sequence 3159, Ap
C 126	14.2	71.0	6464	4	US-09-221-017B-168	Sequence 168, App	C 199	13.6	68.0	593	3	US-08-961-083-107	Sequence 107, App
C 127	14.2	71.0	7286	3	US-09-331-581-3	Sequence 3, Appli	C 200	13.6	68.0	593	3	US-09-536-784-107	Sequence 107, App
C 128	14.2	71.0	7938	4	US-09-331-581-14	Sequence 14, Appl	C 201	13.6	68.0	642	4	US-08-833-381-1086	Sequence 1086, Ap
C 129	14.2	71.0	11340	4	US-08-961-527-147	Sequence 147, App	C 202	13.6	68.0	680	4	US-09-227-357-144	Sequence 144, App
C 130	14.2	71.0	20710	1	US-08-420-235B-1	Sequence 1, Appli	C 203	13.6	68.0	751	4	US-09-621-976-3052	Sequence 3052, Ap
C 131	14.2	71.0	20710	1	US-08-793-624-1	Sequence 1, Appli	C 204	13.6	68.0	876	2	US-08-928-28A-1	Sequence 1, Appli
C 132	14.2	71.0	20710	5	US-08-931-581-1	Sequence 1, Appli	C 205	13.6	68.0	975	3	US-09-365-150-3	Sequence 3, Appli
C 133	14.2	71.0	35100	3	US-08-770-379-18	Sequence 18, Appl	C 206	13.6	68.0	975	3	US-09-365-150-4	Sequence 4, Appli
C 134	14.2	71.0	35100	3	US-08-757-669A-18	Sequence 18, Appl	C 207	13.6	68.0	1158	4	US-09-491-577-41	Sequence 41, Appl
C 135	14.2	71.0	35100	4	US-09-230-371A-18	Sequence 18, Appl	C 208	13.6	68.0	1197	3	US-09-383-586-5	Sequence 5, Appli
C 136	14.2	71.0	36063	4	US-08-311-731A-140	Sequence 140, App	C 209	13.6	68.0	1197	4	US-09-494-921-1	Sequence 1, Appli
C 137	14.2	71.0	36241	4	US-08-311-731A-134	Sequence 134, App	C 210	13.6	68.0	1197	4	US-09-494-921-28	Sequence 28, Appli
C 138	14.2	71.0	42157	4	US-08-311-731A-126	Sequence 126, App	C 211	13.6	68.0	1197	4	US-09-494-921-29	Sequence 29, Appli
C 139	14.2	71.0	580073	4	US-08-545-528D-1	Sequence 1, Appli	C 212	13.6	68.0	1202	4	US-09-494-921-30	Sequence 30, Appli
C 140	14.2	71.0	580073	4	US-08-545-528D-1	Sequence 1, Appli	C 213	13.6	68.0	1202	4	US-09-494-921-3	Sequence 3, Appli
C 141	14.2	71.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli	C 214	13.6	68.0	1209	4	US-09-489-039A-344	Sequence 344, App
C 142	14.2	71.0	435	4	US-09-328-352-1934	Sequence 1934, Ap	C 215	13.6	68.0	1216	3	US-08-838-111A-62	Sequence 62, Appli
C 143	14.2	71.0	2079	3	US-09-489-847-25	Sequence 25, Appl	C 216	13.6	68.0	1247	4	US-09-740-035-1	Sequence 1, Appli
C 144	13.8	69.0	279	3	US-09-328-111-613	Sequence 613, App	C 217	13.6	68.0	1260	3	US-09-008-979A-3	Sequence 3, Appli
C 145	13.8	69.0	426	4	US-08-470-179-69	Sequence 69, Appl	C 218	13.6	68.0	1260	3	US-09-460-618-3	Sequence 3, Appli
C 146	13.8	69.0	466	4	US-09-621-976-11011	Sequence 11011, A	C 219	13.6	68.0	1260	4	US-09-310-235B-3	Sequence 3, Appli
C 147	13.8	69.0	486	4	US-09-543-681A-2354	Sequence 2354, Ap	C 220	13.6	68.0	1263	4	US-09-252-981A-10082	Sequence 10082, A
C 148	13.8	69.0	633	3	US-09-352-990-15	Sequence 15, Appl	C 221	13.6	68.0	1302	4	US-09-170-466D-47	Sequence 47, Appl
C 149	13.8	69.0	638	4	US-09-669-751-66	Sequence 66, Appl	C 222	13.6	68.0	1302	4	US-09-170-466D-189	Sequence 189, App
C 150	13.8	69.0	741	4	US-09-489-039A-3825	Sequence 3825, Ap	C 223	13.6	68.0	1320	2	US-08-641-038A-1	Sequence 1, Appli
C 151	13.8	69.0	766	2	US-08-763-121-2	Sequence 2, Appli	C 224	13.6	68.0	1320	2	US-09-059-178-1	Sequence 1, Appli
C 152	13.8	69.0	766	4	US-09-216-066-2	Sequence 2, Appli	C 225	13.6	68.0	1329	4	US-09-328-352-2751	Sequence 2751, Ap
C 153	13.8	69.0	1107	1	US-08-844-055-1	Sequence 1, Appli	C 226	13.6	68.0	1332	4	US-09-134-001C-311	Sequence 311, App
C 154	13.8	69.0	1107	3	US-09-006-849-1	Sequence 1, Appli	C 227	13.6	68.0	1335	4	US-09-540-226-1425	Sequence 1425, Ap
C 155	13.8	69.0	1119	4	US-09-252-991A-8546	Sequence 8546, Ap	C 228	13.6	68.0	1359	4	US-09-489-039A-5387	Sequence 5387, Ap
C 156	13.8	69.0	1175	2	US-08-773-423-6	Sequence 6, Appli	C 229	13.6	68.0	1359	2	US-08-642-541-1	Sequence 1, Appli
C 157	13.8	69.0	1290	3	US-08-906-744A-1	Sequence 1, Appli	C 230	13.6	68.0	1369	2	US-08-642-541-3	Sequence 3, Appli
C 158	13.8	69.0	1290	3	US-09-093-134-1	Sequence 1, Appli	C 231	13.6	68.0	1369	3	US-09-260-889-1	Sequence 1, Appli
C 159	13.8	69.0	1458	4	US-09-482-273-14	Sequence 14, Appl	C 232	13.6	68.0	1369	3	US-09-260-889-3	Sequence 3, Appli
C 160	13.8	69.0	1464	4	US-09-328-352-3361	Sequence 3361, Ap	C 233	13.6	68.0	1369	4	US-09-347-878-11	Sequence 11, Appli
C 161	13.8	69.0	1550	3	US-08-943-731-70	Sequence 70, Appl	C 234	13.6	68.0	1369	4	US-09-479-275-6	Sequence 6, Appli
C 162	13.8	69.0	1569	4	US-09-328-352-2714	Sequence 2714, Ap	C 235	13.6	68.0	1380	4	US-09-134-001C-2629	Sequence 2629, Ap
C 163	13.8	69.0	1740	4	US-09-489-039A-5163	Sequence 5163, Ap	C 236	13.6	68.0	1457	4	US-09-419-679-9	Sequence 9, Appli
C 164	13.8	69.0	2029	4	US-09-136-574A-46	Sequence 46, Appl	C 237	13.6	68.0	1479	4	US-09-252-991A-6291	Sequence 6291, Ap
C 165	13.8	69.0	2184	4	US-09-543-681A-2518	Sequence 2518, Ap	C 238	13.6	68.0	1512	4	US-09-107-532A-54	Sequence 54, Appl
C 166	13.8	69.0	3311	2	US-08-239-276-10	Sequence 10, Appl	C 239	13.6	68.0	1662	4	US-09-252-991A-6018	Sequence 6018, Ap
C 167	13.8	69.0	3311	2	US-08-468-579B-10	Sequence 10, Appl	C 240	13.6	68.0	1721	4	US-09-197-679A-1	Sequence 1, Appli
C 168	13.8	69.0	3311	3	US-08-468-577B-10	Sequence 10, Appl	C 241	13.6	68.0	1734	4	US-09-506-066E-3	Sequence 3, Appli
C 169	13.8	69.0	3606	4	US-08-956-171E-43	Sequence 43, Appl	C 242	13.6	68.0	1796	1	US-07-816-283-11	Sequence 11, Appli
C 170	13.8	69.0	3613	4	US-08-514-213A-1	Sequence 5, Appli	C 243	13.6	68.0	1796	1	US-08-417-103-11	Sequence 11, Appli
C 171	13.8	69.0	3637	4	US-09-866-921-5	Sequence 5, Appli	C 244	13.6	68.0	1855	4	US-09-403-752A-19	Sequence 19, Appli
C 172	13.8	69.0	3797	1	US-07-915-203-1	Sequence 1, Appli	C 245	13.6	68.0	2000	4	US-07-923-739-1	Sequence 1, Appli
C 173	13.8	69.0	3797	1	US-08-872-887-1	Sequence 1, Appli	C 246	13.6	68.0	2002	2	US-08-715-202A-3	Sequence 3, Appli

C 247	13.6	68.0	2002	4	US-09-328-775-3	Sequence 3, Appl1	320	13.2	66.0	185	4	US-09-313-294A-4007	Sequence 4007, Ap
C 248	13.6	68.0	2002	4	US-09-994-177-3	Sequence 3, Appl1	C 321	13.2	66.0	238	1	US-08-435-684A-56	Sequence 56, Appl
C 249	13.6	68.0	2045	4	US-08-753-750B-5	Sequence 5, Appl1	C 322	13.2	66.0	238	2	US-08-934-877A-56	Sequence 56, Appl
C 250	13.6	68.0	2211	4	US-09-016-434-1150	Sequence 1150, Ap	C 323	13.2	66.0	238	3	US-08-871-678B-56	Sequence 56, Appl
C 251	13.6	68.0	2238	2	US-08-919-624-2	Sequence 2, Appl1	C 324	13.2	66.0	289	4	US-09-313-294A-1390	Sequence 1390, Ap
C 252	13.6	68.0	2259	1	US-08-145-681-5	Sequence 5, Appl1	C 325	13.2	66.0	306	4	US-09-543-681A-121	Sequence 121, App
C 253	13.6	68.0	2259	1	US-08-453-703-5	Sequence 5, Appl1	C 326	13.2	66.0	327	3	US-08-766-355-9	Sequence 9, Appl1
C 254	13.6	68.0	2259	2	US-08-456-106-5	Sequence 5, Appl1	C 327	13.2	66.0	327	3	US-09-003-198A-9	Sequence 9, Appl1
C 255	13.6	68.0	2259	3	US-08-456-108-5	Sequence 5, Appl1	C 328	13.2	66.0	327	3	US-09-428-805-9	Sequence 9, Appl1
C 256	13.6	68.0	2259	3	US-09-265-577-5	Sequence 5, Appl1	C 329	13.2	66.0	352	3	US-08-961-083-127	Sequence 127, App
C 257	13.6	68.0	2259	4	US-09-633-739-5	Sequence 5, Appl1	C 330	13.2	66.0	352	4	US-09-536-784-127	Sequence 127, App
C 258	13.6	68.0	2278	4	US-09-276-438-10	Sequence 10, Appl1	C 331	13.2	66.0	384	4	US-09-543-681A-1889	Sequence 1889, Ap
C 259	13.6	68.0	2587	2	US-08-326-286-3	Sequence 3, Appl1	C 332	13.2	66.0	396	4	US-09-621-976-1322	Sequence 8322, Ap
C 260	13.6	68.0	2817	4	US-08-956-171E-295	Sequence 295, App	C 333	13.2	66.0	430	4	US-09-621-976-1254	Sequence 8254, Ap
C 261	13.6	68.0	2896	2	US-08-709-923-1	Sequence 1, Appl1	C 334	13.2	66.0	445	4	US-09-621-976-15305	Sequence 15305, A
C 262	13.6	68.0	3375	1	US-08-381-931B-1	Sequence 1, Appl1	C 335	13.2	66.0	482	4	US-09-621-976-115	Sequence 315, App
C 263	13.6	68.0	3466	4	US-09-782-906-1	Sequence 1, Appl1	C 336	13.2	66.0	490	4	US-09-003-198A-19	Sequence 19, Appl1
C 264	13.6	68.0	3784	1	US-07-623-033-1	Sequence 1, Appl1	C 337	13.2	66.0	495	4	US-09-543-681A-2833	Sequence 2833, Ap
C 265	13.6	68.0	3993	4	US-09-543-681A-3885	Sequence 3885, Ap	C 338	13.2	66.0	507	3	US-08-766-355-10	Sequence 10, Appl1
C 266	13.6	68.0	4081	3	US-08-999-774A-1	Sequence 1, Appl1	C 339	13.2	66.0	507	4	US-09-003-198A-10	Sequence 10, Appl1
C 267	13.6	68.0	4617	4	US-08-930-055A-1	Sequence 1, Appl1	C 340	13.2	66.0	507	4	US-09-428-805-10	Sequence 10, Appl1
C 268	13.6	68.0	4761	4	US-09-543-681A-1157	Sequence 1157, Ap	C 341	13.2	66.0	510	4	US-09-134-001C-1698	Sequence 1698, Ap
C 269	13.6	68.0	5024	1	US-08-820-812-7	Sequence 7, Appl1	C 342	13.2	66.0	522	4	US-09-598-401C-61	Sequence 61, Appl
C 270	13.6	68.0	5024	1	US-08-920-827-7	Sequence 7, Appl1	C 343	13.2	66.0	537	4	US-09-621-976-10463	Sequence 10463, A
C 271	13.6	68.0	5024	1	US-08-921-177-7	Sequence 7, Appl1	C 344	13.2	66.0	538	4	US-09-621-976-1523	Sequence 2523, Ap
C 272	13.6	68.0	5024	1	US-08-362-577C-7	Sequence 7, Appl1	C 345	13.2	66.0	600	4	US-09-621-976-2524	Sequence 2524, Ap
C 273	13.6	68.0	5024	2	US-08-920-828-7	Sequence 7, Appl1	C 346	13.2	66.0	640	3	US-09-328-111-792	Sequence 792, App
C 274	13.6	68.0	5109	4	US-08-930-055A-2	Sequence 2, Appl1	C 347	13.2	66.0	651	4	US-09-134-001C-781	Sequence 781, App
C 275	13.6	68.0	5455	4	US-08-342-930-1	Sequence 1, Appl1	C 348	13.2	66.0	688	4	US-09-972-800A-70	Sequence 70, Appl
C 276	13.6	68.0	6102	4	US-09-976-594-682	Sequence 682, App	C 349	13.2	66.0	767	4	US-09-921-017B-28	Sequence 28, Appl
C 277	13.6	68.0	6102	4	US-08-961-527-167	Sequence 167, App	C 350	13.2	66.0	813	4	US-09-872-733A-19	Sequence 19, Appl1
C 278	13.6	68.0	29555	4	US-08-956-171E-206	Sequence 206, App	C 351	13.2	66.0	846	1	US-07-672-304-2	Sequence 2, Appl1
C 279	13.6	68.0	29555	4	US-08-956-171E-306	Sequence 206, App	C 352	13.2	66.0	870	4	US-09-328-352-1901	Sequence 1901, Ap
C 280	13.6	68.0	36800	4	US-08-311-731A-139	Sequence 139, App	C 353	13.2	66.0	876	4	US-09-134-000C-860	Sequence 860, App
C 281	13.6	68.0	37769	4	US-08-311-731A-23	Sequence 23, Appl	C 354	13.2	66.0	885	4	US-09-543-681A-1746	Sequence 1746, App
C 282	13.6	68.0	56516	2	US-08-996-306-1	Sequence 1, Appl1	C 355	13.2	66.0	887	4	US-09-247-155-150	Sequence 150, App
C 283	13.6	68.0	56516	3	US-09-338-907-1	Sequence 1, Appl1	C 356	13.2	66.0	899	4	US-09-221-017B-1060	Sequence 1060, App
C 284	13.6	68.0	56516	4	US-09-218-207-1	Sequence 1, Appl1	C 357	13.2	66.0	914	4	US-09-679-395A-114	Sequence 114, App
C 285	13.6	68.0	56520	3	US-09-338-907-179	Sequence 179, App	C 358	13.2	66.0	921	4	US-09-489-039A-1033	Sequence 1033, Ap
C 286	13.6	68.0	56520	4	US-09-218-207-179	Sequence 179, App	C 359	13.2	66.0	923	4	US-09-221-017B-252	Sequence 252, App
C 287	13.6	68.0	66986	4	US-09-596-002-29	Sequence 29, Appl	C 360	13.2	66.0	933	4	US-09-107-532A-558	Sequence 558, App
C 288	13.6	68.0	392000	4	US-10-027-983-11	Sequence 11, Appl1	C 361	13.2	66.0	969	4	US-09-543-681A-2787	Sequence 2787, App
C 289	13.6	68.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl1	C 362	13.2	66.0	969	4	US-09-543-681A-3696	Sequence 3696, Ap
C 290	13.6	68.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl1	C 363	13.2	66.0	1002	4	US-09-641-638-883	Sequence 583, App
C 291	13.4	67.0	266	4	US-09-313-294A-1855	Sequence 1855, Ap	C 364	13.2	66.0	1007	4	US-09-620-312D-817	Sequence 817, App
C 292	13.4	67.0	311	4	US-09-963-137-15	Sequence 15, Appl	C 365	13.2	66.0	1021	4	US-08-936-11E-11	Sequence 11, Appl
C 293	13.4	67.0	431	4	US-09-621-976-16186	Sequence 16186, A	C 366	13.2	66.0	1035	4	US-09-134-001C-687	Sequence 687, App
C 294	13.4	67.0	442	4	US-09-071-035-443	Sequence 443, App	C 367	13.2	66.0	1092	4	US-09-797-464A-10	Sequence 10, Appl
C 295	13.4	67.0	453	4	US-08-817-441-64	Sequence 64, Appl	C 368	13.2	66.0	1113	4	US-09-540-236A-138	Sequence 338, App
C 296	13.4	67.0	606	4	US-09-071-035-441	Sequence 441, App	C 369	13.2	66.0	1116	3	US-08-916-443A-16	Sequence 16, Appl
C 297	13.4	67.0	638	2	US-08-222-721-16	Sequence 16, Appl	C 370	13.2	66.0	1142	4	US-08-936-165A-92	Sequence 92, Appl
C 298	13.4	67.0	638	2	US-08-470-925-16	Sequence 16, Appl	C 371	13.2	66.0	1233	3	US-08-853-839-1	Sequence 1, Appl1
C 299	13.4	67.0	638	2	US-08-471-613-16	Sequence 16, Appl	C 372	13.2	66.0	1239	4	US-09-252-991A-12224	Sequence 12224, A
C 300	13.4	67.0	638	5	PCT-US93-10443-16	Sequence 16, Appl	C 373	13.2	66.0	1242	4	US-09-543-681A-1921	Sequence 1921, App
C 301	13.4	67.0	1123	4	US-09-134-000C-1299	Sequence 1299, Ap	C 374	13.2	66.0	1248	2	US-08-897-340-5	Sequence 5, Appl1
C 302	13.4	67.0	1290	4	US-08-956-171E-427	Sequence 427, App	C 375	13.2	66.0	1248	2	US-08-897-340-5	Sequence 5, Appl1
C 303	13.4	67.0	1473	3	US-08-907-740-6	Sequence 6, Appl1	C 376	13.2	66.0	1291	1	US-07-952-755-2	Sequence 2, Appl1
C 304	13.4	67.0	2616	4	US-09-107-532A-2573	Sequence 2573, Ap	C 377	13.2	66.0	1291	1	US-08-443-679-2	Sequence 215, App
C 305	13.4	67.0	2715	4	US-09-543-681A-513	Sequence 513, App	C 378	13.2	66.0	1308	4	US-09-205-258-115	Sequence 80, Appl1
C 306	13.4	67.0	3301	4	US-09-539-333D-213	Sequence 213, App	C 379	13.2	66.0	1360	2	US-09-082-649B-80	Sequence 2, Appl1
C 307	13.4	67.0	3379	4	US-09-275-608-1	Sequence 1, Appl1	C 380	13.2	66.0	1378	2	US-08-776-585-2	Sequence 2, Appl1
C 308	13.4	67.0	4695	4	US-09-620-312D-379	Sequence 379, App	C 381	13.2	66.0	1400	2	US-08-776-585-1	Sequence 1, Appl1
C 309	13.4	67.0	6543	4	US-09-963-137-131	Sequence 131, App	C 382	13.2	66.0	1404	4	US-09-489-039A-4718	Sequence 4718, Ap
C 310	13.4	67.0	8608	4	US-09-620-312D-333	Sequence 333, App	C 383	13.2	66.0	1414	4	US-09-404-268B-9	Sequence 9, Appl1
C 311	13.4	67.0	24979	2	US-08-147-777-3	Sequence 3, Appl1	C 384	13.2	66.0	1429	2	US-08-244-205-6	Sequence 6, Appl1
C 312	13.4	67.0	24979	5	PCT-US93-03985-3	Sequence 3, Appl1	C 385	13.2	66.0	1429	5	PCT-US92-10284-6	Sequence 8, Appl1
C 313	13.4	67.0	40138	3	US-09-090-793-112	Sequence 12, Appl	C 386	13.2	66.0	1429	5	PCT-US92-10284-6	Sequence 8, Appl1
C 314	13.4	67.0	40138	4	US-09-231-899-12	Sequence 12, Appl	C 387	13.2	66.0	1429	5	PCT-US92-10284-6	Sequence 8, Appl1
C 315	13.4	67.0	40138	4	US-08-947-823-1	Sequence 1, Appl1	C 388	13.2	66.0	1455	4	US-09-489-039A-4800	Sequence 4800, Ap
C 316	13.4	67.0	51952	3	US-08-947-823-1	Sequence 1, Appl1	C 389	13.2	66.0	1457	4	US-09-673-355A-554	Sequence 554, App
C 317	13.4	67.0	536165	4	US-09-214-808-1	Sequence 3, Appl1	C 390	13.2	66.0	1479	4	US-09-489-039A-1847	Sequence 1847, App
C 318	13.4	67.0	786431	4	US-09-751-389-3	Sequence 58, Appl1	C 391	13.2	66.0	1482	3	US-08-996-441B-69	Sequence 69, Appl
C 319	13.2	66.0	20	4	US-09-972-800A-58	Sequence 58, Appl1	C 392	13.2	66.0	1482	3	US-08-993-722A-69	Sequence 69, Appl

C 393	13.2	66.0	1482	3	US-08-993-170A-69	Sequence 69, Appl	C 466	13.2	66.0	1959	3	US-08-993-722A-47	Sequence 47, Appl
C 394	13.2	66.0	1482	3	US-08-993-775B-69	Sequence 69, Appl	C 467	13.2	66.0	1959	3	US-08-993-722A-49	Sequence 49, Appl
C 395	13.2	66.0	1482	3	US-09-427-770-69	Sequence 69, Appl	C 468	13.2	66.0	1959	3	US-08-993-722A-53	Sequence 53, Appl
C 396	13.2	66.0	1482	4	US-09-427-769-69	Sequence 69, Appl	C 469	13.2	66.0	1959	3	US-08-993-722A-63	Sequence 63, Appl
C 397	13.2	66.0	1486	4	US-07-672-304-5	Sequence 5, Appl	C 470	13.2	66.0	1959	3	US-08-993-722A-65	Sequence 65, Appl
C 398	13.2	66.0	1483	4	US-09-227-357-26	Sequence 26, Appl	C 471	13.2	66.0	1959	3	US-08-993-722A-67	Sequence 67, Appl
C 399	13.2	66.0	1600	3	US-08-917-299-36	Sequence 36, Appl	C 472	13.2	66.0	1959	3	US-08-993-722A-97	Sequence 97, Appl
C 400	13.2	66.0	1600	3	US-09-422-629-36	Sequence 36, Appl	C 473	13.2	66.0	1959	3	US-08-993-170A-1	Sequence 1, Appl
C 401	13.2	66.0	1600	4	US-09-730-763-36	Sequence 36, Appl	C 474	13.2	66.0	1959	3	US-08-993-170A-3	Sequence 3, Appl
C 402	13.2	66.0	1600	4	US-09-429-370-36	Sequence 36, Appl	C 475	13.2	66.0	1959	3	US-08-993-170A-5	Sequence 5, Appl
C 403	13.2	66.0	1644	4	US-09-489-039A-5502	Sequence 5202, Ap	C 476	13.2	66.0	1959	3	US-08-993-170A-7	Sequence 7, Appl
C 404	13.2	66.0	1447	4	US-09-543-681A-2459	Sequence 2459, Ap	C 477	13.2	66.0	1959	3	US-08-993-170A-9	Sequence 9, Appl
C 405	13.2	66.0	1682	4	US-09-620-312D-181	Sequence 181, Ap	C 478	13.2	66.0	1959	3	US-08-993-170A-11	Sequence 11, Appl
C 406	13.2	66.0	1691	3	US-08-917-299-3	Sequence 3, Appl	C 479	13.2	66.0	1959	3	US-08-993-170A-13	Sequence 13, Appl
C 407	13.2	66.0	1691	3	US-09-422-662-3	Sequence 3, Appl	C 480	13.2	66.0	1959	3	US-08-993-170A-27	Sequence 27, Appl
C 408	13.2	66.0	1691	3	US-09-730-763-3	Sequence 3, Appl	C 481	13.2	66.0	1959	3	US-08-993-170A-29	Sequence 29, Appl
C 409	13.2	66.0	1691	4	US-09-230-485-1	Sequence 1, Appl	C 482	13.2	66.0	1959	3	US-08-993-170A-39	Sequence 39, Appl
C 410	13.2	66.0	1738	4	US-09-574-873-9	Sequence 9, Appl	C 483	13.2	66.0	1959	3	US-08-993-170A-41	Sequence 41, Appl
C 411	13.2	66.0	1875	4	US-09-543-681A-1964	Sequence 1964, Ap	C 484	13.2	66.0	1959	3	US-08-993-170A-43	Sequence 43, Appl
C 412	13.2	66.0	1917	4	US-09-107-532A-265	Sequence 265, Ap	C 485	13.2	66.0	1959	3	US-08-993-170A-45	Sequence 45, Appl
C 413	13.2	66.0	1953	1	US-08-315-468-5	Sequence 5, Appl	C 486	13.2	66.0	1959	3	US-08-993-170A-47	Sequence 47, Appl
C 414	13.2	66.0	1953	3	US-07-941-650A-3	Sequence 3, Appl	C 487	13.2	66.0	1959	3	US-08-993-170A-49	Sequence 49, Appl
C 415	13.2	66.0	1956	3	US-08-996-441B-51	Sequence 51, Appl	C 488	13.2	66.0	1959	3	US-08-993-170A-53	Sequence 53, Appl
C 416	13.2	66.0	1956	3	US-08-996-441B-55	Sequence 55, Appl	C 489	13.2	66.0	1959	3	US-08-993-170A-65	Sequence 65, Appl
C 417	13.2	66.0	1956	3	US-08-996-441B-57	Sequence 57, Appl	C 490	13.2	66.0	1959	3	US-08-993-170A-67	Sequence 67, Appl
C 418	13.2	66.0	1956	3	US-08-993-722A-51	Sequence 51, Appl	C 491	13.2	66.0	1959	3	US-08-993-170A-69	Sequence 69, Appl
C 419	13.2	66.0	1956	3	US-08-993-722A-55	Sequence 55, Appl	C 492	13.2	66.0	1959	3	US-08-993-170A-97	Sequence 97, Appl
C 420	13.2	66.0	1956	3	US-08-993-722A-57	Sequence 57, Appl	C 493	13.2	66.0	1959	3	US-08-993-775B-3	Sequence 3, Appl
C 421	13.2	66.0	1956	3	US-08-993-170A-51	Sequence 51, Appl	C 494	13.2	66.0	1959	3	US-08-993-775B-5	Sequence 5, Appl
C 422	13.2	66.0	1956	3	US-08-993-170A-55	Sequence 55, Appl	C 495	13.2	66.0	1959	3	US-08-993-775B-7	Sequence 7, Appl
C 423	13.2	66.0	1956	3	US-08-993-170A-57	Sequence 57, Appl	C 496	13.2	66.0	1959	3	US-08-993-775B-9	Sequence 9, Appl
C 424	13.2	66.0	1956	3	US-08-993-775B-51	Sequence 51, Appl	C 497	13.2	66.0	1959	3	US-08-993-775B-11	Sequence 11, Appl
C 425	13.2	66.0	1956	3	US-08-993-775B-55	Sequence 55, Appl	C 498	13.2	66.0	1959	3	US-08-993-775B-13	Sequence 13, Appl
C 426	13.2	66.0	1956	3	US-08-993-775B-57	Sequence 57, Appl	C 499	13.2	66.0	1959	3	US-08-993-775B-45	Sequence 45, Appl
C 427	13.2	66.0	1956	4	US-09-427-770-51	Sequence 51, Appl	C 500	13.2	66.0	1959	3	US-08-993-775B-47	Sequence 47, Appl
C 428	13.2	66.0	1956	4	US-09-427-770-55	Sequence 55, Appl	C 501	13.2	66.0	1959	3	US-08-993-775B-29	Sequence 29, Appl
C 429	13.2	66.0	1956	4	US-09-427-770-57	Sequence 57, Appl	C 502	13.2	66.0	1959	3	US-08-993-775B-39	Sequence 39, Appl
C 430	13.2	66.0	1956	4	US-09-427-769-51	Sequence 51, Appl	C 503	13.2	66.0	1959	3	US-08-993-775B-43	Sequence 43, Appl
C 431	13.2	66.0	1956	4	US-09-427-769-55	Sequence 55, Appl	C 504	13.2	66.0	1959	3	US-08-993-775B-45	Sequence 45, Appl
C 432	13.2	66.0	1956	4	US-09-427-769-57	Sequence 57, Appl	C 505	13.2	66.0	1959	3	US-08-993-775B-47	Sequence 47, Appl
C 433	13.2	66.0	1959	3	US-08-996-441B-1	Sequence 1, Appl	C 506	13.2	66.0	1959	3	US-08-993-775B-49	Sequence 49, Appl
C 434	13.2	66.0	1959	3	US-08-996-441B-3	Sequence 3, Appl	C 507	13.2	66.0	1959	3	US-08-993-775B-53	Sequence 53, Appl
C 435	13.2	66.0	1959	3	US-08-996-441B-5	Sequence 5, Appl	C 508	13.2	66.0	1959	3	US-08-993-775B-65	Sequence 65, Appl
C 436	13.2	66.0	1959	3	US-08-996-441B-7	Sequence 7, Appl	C 509	13.2	66.0	1959	3	US-08-993-775B-67	Sequence 67, Appl
C 437	13.2	66.0	1959	3	US-08-996-441B-11	Sequence 11, Appl	C 510	13.2	66.0	1959	3	US-08-993-775B-69	Sequence 69, Appl
C 438	13.2	66.0	1959	3	US-08-996-441B-13	Sequence 13, Appl	C 511	13.2	66.0	1959	3	US-08-993-775B-7	Sequence 7, Appl
C 439	13.2	66.0	1959	3	US-08-996-441B-27	Sequence 27, Appl	C 512	13.2	66.0	1959	3	US-09-377-466B-1	Sequence 1, Appl
C 440	13.2	66.0	1959	3	US-08-996-441B-29	Sequence 29, Appl	C 513	13.2	66.0	1959	4	US-09-377-466B-3	Sequence 3, Appl
C 441	13.2	66.0	1959	3	US-08-996-441B-39	Sequence 39, Appl	C 515	13.2	66.0	1959	4	US-09-427-770-1	Sequence 1, Appl
C 442	13.2	66.0	1959	3	US-08-996-441B-41	Sequence 41, Appl	C 516	13.2	66.0	1959	4	US-09-427-770-3	Sequence 3, Appl
C 443	13.2	66.0	1959	3	US-08-996-441B-45	Sequence 45, Appl	C 517	13.2	66.0	1959	4	US-09-427-770-5	Sequence 5, Appl
C 444	13.2	66.0	1959	3	US-08-996-441B-47	Sequence 47, Appl	C 518	13.2	66.0	1959	4	US-09-427-770-7	Sequence 7, Appl
C 445	13.2	66.0	1959	3	US-08-996-441B-49	Sequence 49, Appl	C 519	13.2	66.0	1959	4	US-09-427-770-9	Sequence 9, Appl
C 446	13.2	66.0	1959	3	US-08-996-441B-53	Sequence 53, Appl	C 520	13.2	66.0	1959	4	US-09-427-770-11	Sequence 11, Appl
C 447	13.2	66.0	1959	3	US-08-996-441B-55	Sequence 55, Appl	C 521	13.2	66.0	1959	4	US-09-427-770-13	Sequence 13, Appl
C 448	13.2	66.0	1959	3	US-08-996-441B-57	Sequence 57, Appl	C 522	13.2	66.0	1959	4	US-09-427-770-17	Sequence 17, Appl
C 449	13.2	66.0	1959	3	US-08-996-441B-65	Sequence 65, Appl	C 523	13.2	66.0	1959	4	US-09-427-770-29	Sequence 29, Appl
C 450	13.2	66.0	1959	3	US-08-996-441B-67	Sequence 67, Appl	C 524	13.2	66.0	1959	4	US-09-427-770-39	Sequence 39, Appl
C 451	13.2	66.0	1959	3	US-08-996-441B-69	Sequence 69, Appl	C 525	13.2	66.0	1959	4	US-09-427-770-41	Sequence 41, Appl
C 452	13.2	66.0	1959	3	US-08-996-441B-7	Sequence 7, Appl	C 526	13.2	66.0	1959	4	US-09-427-770-43	Sequence 43, Appl
C 453	13.2	66.0	1959	3	US-08-993-722A-1	Sequence 1, Appl	C 527	13.2	66.0	1959	4	US-09-427-770-45	Sequence 45, Appl
C 454	13.2	66.0	1959	3	US-08-993-722A-3	Sequence 3, Appl	C 528	13.2	66.0	1959	4	US-09-427-770-47	Sequence 47, Appl
C 455	13.2	66.0	1959	3	US-08-993-722A-5	Sequence 5, Appl	C 529	13.2	66.0	1959	4	US-09-427-770-49	Sequence 49, Appl
C 456	13.2	66.0	1959	3	US-08-993-722A-7	Sequence 7, Appl	C 530	13.2	66.0	1959	4	US-09-427-770-53	Sequence 53, Appl
C 457	13.2	66.0	1959	3	US-08-993-722A-9	Sequence 9, Appl	C 531	13.2	66.0	1959	4	US-09-427-770-63	Sequence 63, Appl
C 458	13.2	66.0	1959	3	US-08-993-722A-11	Sequence 11, Appl	C 532	13.2	66.0	1959	4	US-09-427-770-65	Sequence 65, Appl
C 459	13.2	66.0	1959	3	US-08-993-722A-13	Sequence 13, Appl	C 533	13.2	66.0	1959	4	US-09-427-770-67	Sequence 67, Appl
C 460	13.2	66.0	1959	3	US-08-993-722A-27	Sequence 27, Appl	C 534	13.2	66.0	1959	4	US-09-427-769-1	Sequence 1, Appl
C 461	13.2	66.0	1959	3	US-08-993-722A-29	Sequence 29, Appl	C 535	13.2	66.0	1959	4	US-09-427-769-3	Sequence 3, Appl
C 462	13.2	66.0	1959	3	US-08-993-722A-39	Sequence 39, Appl	C 537	13.2	66.0	1959	4	US-09-427-769-5	Sequence 5, Appl
C 463	13.2	66.0	1959	3	US-08-993-722A-41	Sequence 41, Appl	C 538	13.2	66.0	1959	4	US-09-427-769-7	Sequence 7, Appl
C 464	13.2	66.0	1959	3	US-08-993-722A-43	Sequence 43, Appl							
C 465	13.2	66.0	1959	3	US-08-993-722A-45	Sequence 45, Appl							



C 539	13.2	66.0	1959	4	US-09-427-769-9	Sequence 9, Appli	612	13.2	66.0	4505	5	PCT-US95-12642-13	Sequence 13, Appli
C 540	13.2	66.0	1959	4	US-09-427-769-11	Sequence 11, Appli	613	13.2	66.0	4526	4	US-09-424-283-7	Sequence 7, Appli
C 541	13.2	66.0	1959	4	US-09-427-769-11	Sequence 13, Appli	614	13.2	66.0	4466	4	US-09-485-286-14	Sequence 14, Appli
C 542	13.2	66.0	1959	4	US-09-427-769-27	Sequence 27, Appli	615	13.2	66.0	4700	2	US-08-928-632-16	Sequence 16, Appli
C 543	13.2	66.0	1959	4	US-09-427-769-29	Sequence 29, Appli	616	13.2	66.0	4700	2	US-08-928-632-16	Sequence 16, Appli
C 544	13.2	66.0	1959	4	US-09-427-769-39	Sequence 39, Appli	617	13.2	66.0	4771	2	US-08-866-650-2	Sequence 2, Appli
C 545	13.2	66.0	1959	4	US-09-427-769-41	Sequence 41, Appli	618	13.2	66.0	4771	2	US-09-021-287-2	Sequence 2, Appli
C 546	13.2	66.0	1959	4	US-09-427-769-43	Sequence 43, Appli	619	13.2	66.0	4771	3	US-09-240-473-2	Sequence 2, Appli
C 547	13.2	66.0	1959	4	US-09-427-769-45	Sequence 45, Appli	620	13.2	66.0	4864	4	US-09-340-798A-1	Sequence 1, Appli
C 548	13.2	66.0	1959	4	US-09-427-769-47	Sequence 47, Appli	621	13.2	66.0	4918	1	US-08-173-053-7	Sequence 7, Appli
C 549	13.2	66.0	1959	4	US-09-427-769-49	Sequence 49, Appli	622	13.2	66.0	4928	1	US-08-345-913-1	Sequence 1, Appli
C 550	13.2	66.0	1959	4	US-09-427-769-53	Sequence 53, Appli	623	13.2	66.0	4928	3	US-08-818-562-1	Sequence 1, Appli
C 551	13.2	66.0	1959	4	US-09-427-769-63	Sequence 63, Appli	624	13.2	66.0	4928	3	US-08-818-562-1	Sequence 1, Appli
C 552	13.2	66.0	1959	4	US-09-427-769-65	Sequence 65, Appli	625	13.2	66.0	4965	2	US-08-628-445-1	Sequence 1, Appli
C 553	13.2	66.0	1959	4	US-09-427-769-67	Sequence 67, Appli	626	13.2	66.0	4965	2	PCT-US94-05663-1	Sequence 1, Appli
C 554	13.2	66.0	1959	4	US-09-427-769-97	Sequence 97, Appli	627	13.2	66.0	5053	4	US-09-311-784A-35	Sequence 35, Appli
C 555	13.2	66.0	1970	3	US-08-687-590-56	Sequence 56, Appli	628	13.2	66.0	5183	3	US-09-039-555B-18	Sequence 18, Appli
C 556	13.2	66.0	1970	6	US-08-687-590-56	Sequence 56, Appli	629	13.2	66.0	5203	4	US-09-976-594-218	Sequence 218, App
C 557	13.2	66.0	1978	4	US-09-525-223A-8	Sequence 8, Appli	630	13.2	66.0	5215	4	US-09-173-053-8	Sequence 8, Appli
C 558	13.2	66.0	1981	4	US-09-647-390-17	Sequence 17, Appli	631	13.2	66.0	5243	2	US-08-414-335-2	Sequence 2, Appli
C 559	13.2	66.0	2091	4	US-09-107-532A-276	Sequence 276, App	632	13.2	66.0	5502	4	US-09-702-705-785	Sequence 785, App
C 560	13.2	66.0	2115	1	US-08-329-681A-3	Sequence 3, Appli	633	13.2	66.0	5502	4	US-09-736-457-785	Sequence 785, App
C 561	13.2	66.0	2280	3	US-08-996-441B-102	Sequence 102, App	634	13.2	66.0	5502	4	US-09-614-124B-785	Sequence 785, App
C 562	13.2	66.0	2280	3	US-08-993-722A-102	Sequence 102, App	635	13.2	66.0	5502	4	US-09-614-124B-785	Sequence 785, App
C 563	13.2	66.0	2280	3	US-08-993-170A-102	Sequence 102, App	636	13.2	66.0	5502	4	US-09-614-124B-785	Sequence 785, App
C 564	13.2	66.0	2280	3	US-08-993-775B-102	Sequence 102, App	637	13.2	66.0	5592	4	US-09-589-184-785	Sequence 785, App
C 565	13.2	66.0	2280	4	US-09-427-770-102	Sequence 102, App	638	13.2	66.0	5592	4	US-09-495-797-17	Sequence 37, Appli
C 566	13.2	66.0	2280	4	US-09-427-769-102	Sequence 102, App	639	13.2	66.0	5676	2	US-08-663-998-3	Sequence 3, Appli
C 567	13.2	66.0	2392	4	US-09-797-464A-1	Sequence 1, Appli	640	13.2	66.0	5682	2	US-08-663-998-4	Sequence 4, Appli
C 568	13.2	66.0	2430	5	PCT-US92-00040-1	Sequence 1, Appli	641	13.2	66.0	5838	2	US-08-850-860-3	Sequence 3, Appli
C 569	13.2	66.0	2479	4	US-09-050-863-4	Sequence 4, Appli	642	13.2	66.0	5838	2	US-08-944-916-3	Sequence 3, Appli
C 570	13.2	66.0	2479	4	US-09-359-081-4	Sequence 4, Appli	643	13.2	66.0	5838	2	US-08-814-877-3	Sequence 3, Appli
C 571	13.2	66.0	2511	4	US-09-620-312D-559	Sequence 559, App	644	13.2	66.0	5845	4	US-09-272-432A-3	Sequence 3, Appli
C 572	13.2	66.0	2622	4	US-09-543-681A-5143	Sequence 5143, Ap	645	13.2	66.0	5899	4	US-09-173-053-1	Sequence 1, Appli
C 573	13.2	66.0	2624	4	US-09-370-838-19	Sequence 19, Appli	646	13.2	66.0	5900	2	US-09-173-053-2	Sequence 2, Appli
C 574	13.2	66.0	2632	2	US-08-899-324-32	Sequence 32, Appli	647	13.2	66.0	5902	4	US-08-663-998-1	Sequence 1, Appli
C 575	13.2	66.0	2632	2	US-08-329-892B-32	Sequence 32, Appli	648	13.2	66.0	5926	4	US-09-381-989-4	Sequence 4, Appli
C 576	13.2	66.0	2885	4	US-10-164-595-39	Sequence 39, Appli	649	13.2	66.0	5926	4	US-09-027-169-3	Sequence 3, Appli
C 577	13.2	66.0	3224	4	US-08-965-728A-2	Sequence 39, Appli	650	13.2	66.0	5952	4	US-09-027-169-4	Sequence 4, Appli
C 578	13.2	66.0	3311	4	US-10-164-595-11	Sequence 11, Appli	651	13.2	66.0	6044	2	US-08-663-998-2	Sequence 2, Appli
C 579	13.2	66.0	3356	1	US-08-246-403A-1	Sequence 1, Appli	652	13.2	66.0	6044	5	US-08-316-950-18	Sequence 18, Appli
C 580	13.2	66.0	3356	1	US-08-246-403A-4	Sequence 4, Appli	653	13.2	66.0	6044	5	PCT-US95-12642-18	Sequence 18, Appli
C 581	13.2	66.0	3375	4	US-09-309-572-1	Sequence 1, Appli	654	13.2	66.0	6047	5	US-08-316-950-12	Sequence 12, Appli
C 582	13.2	66.0	3375	4	US-09-718-096-1	Sequence 1, Appli	655	13.2	66.0	6171	5	PCT-US95-12642-12	Sequence 12, Appli
C 583	13.2	66.0	3418	2	US-08-944-916-12	Sequence 12, Appli	656	13.2	66.0	6171	5	US-08-316-950-12	Sequence 12, Appli
C 584	13.2	66.0	3424	4	US-09-553-690-3	Sequence 3, Appli	657	13.2	66.0	6171	5	PCT-US95-12642-17	Sequence 17, Appli
C 585	13.2	66.0	3470	4	US-09-976-594-35	Sequence 35, App	658	13.2	66.0	6353	4	US-09-702-705-784	Sequence 784, App
C 586	13.2	66.0	3547	4	US-09-340-798A-43	Sequence 43, Appli	659	13.2	66.0	6353	4	US-09-736-457-784	Sequence 784, App
C 587	13.2	66.0	3610	4	US-09-194-949A-1	Sequence 1, Appli	660	13.2	66.0	6353	4	US-09-614-124B-784	Sequence 784, App
C 588	13.2	66.0	3870	2	US-08-853-659A-26	Sequence 26, Appli	661	13.2	66.0	6353	4	US-09-614-124B-784	Sequence 784, App
C 589	13.2	66.0	3895	4	US-08-961-527-201	Sequence 201, App	662	13.2	66.0	6407	1	US-09-589-184-784	Sequence 784, App
C 590	13.2	66.0	3987	4	US-09-082-649B-83	Sequence 83, Appli	663	13.2	66.0	6407	1	US-07-977-630-1	Sequence 1, Appli
C 591	13.2	66.0	3987	4	US-09-082-649B-84	Sequence 84, Appli	664	13.2	66.0	6407	1	US-07-977-630-2	Sequence 2, Appli
C 592	13.2	66.0	3993	4	US-08-316-950-14	Sequence 14, Appli	665	13.2	66.0	6873	3	US-09-131-028A-8	Sequence 8, Appli
C 593	13.2	66.0	3993	4	PCT-US95-12642-14	Sequence 14, Appli	666	13.2	66.0	6873	3	US-09-131-028A-8	Sequence 8, Appli
C 594	13.2	66.0	4059	5	US-08-564-313-2	Sequence 2, Appli	667	13.2	66.0	6978	4	US-09-872-733A-15	Sequence 15, Appli
C 595	13.2	66.0	4059	5	US-08-564-313-2	Sequence 2, Appli	668	13.2	66.0	6978	4	US-08-975-762-46	Sequence 46, Appli
C 596	13.2	66.0	4066	3	PCT-US94-06069-2	Sequence 2, Appli	669	13.2	66.0	7091	3	US-08-975-762-46	Sequence 46, Appli
C 597	13.2	66.0	4120	1	US-09-298-367B-3	Sequence 3, Appli	670	13.2	66.0	7091	3	US-09-235-028-46	Sequence 46, Appli
C 598	13.2	66.0	4120	1	US-07-977-630-20	Sequence 20, Appli	671	13.2	66.0	7091	4	US-09-106-582-46	Sequence 46, Appli
C 599	13.2	66.0	4120	1	US-07-977-630-21	Sequence 21, Appli	672	13.2	66.0	7091	4	US-09-159-469-46	Sequence 46, Appli
C 600	13.2	66.0	4120	5	US-08-316-950-11	Sequence 11, Appli	673	13.2	66.0	7612	4	US-09-693-542-62	Sequence 46, Appli
C 601	13.2	66.0	4202	4	PCT-US95-12642-11	Sequence 11, Appli	674	13.2	66.0	7612	4	US-09-700-934A-7	Sequence 2, Appli
C 602	13.2	66.0	4261	4	US-09-483-419-2	Sequence 2, Appli	675	13.2	66.0	7676	4	US-07-924-028A-5	Sequence 5, Appli
C 603	13.2	66.0	4296	4	US-09-194-949A-35	Sequence 25, Appli	676	13.2	66.0	7676	4	US-09-036-356-113	Sequence 113, App
C 604	13.2	66.0	4296	5	US-08-316-950-15	Sequence 15, Appli	677	13.2	66.0	7676	4	US-09-072-956-708	Sequence 208, App
C 605	13.2	66.0	4326	3	PCT-US95-12642-15	Sequence 15, Appli	678	13.2	66.0	8031	4	US-09-287-849-9	Sequence 213, App
C 606	13.2	66.0	4338	1	US-08-760-615-7	Sequence 7, Appli	679	13.2	66.0	8031	4	US-09-287-849-9	Sequence 9, Appli
C 607	13.2	66.0	4338	2	US-08-015-986A-1	Sequence 1, Appli	680	13.2	66.0	8031	4	US-09-643-597-254	Sequence 254, App
C 608	13.2	66.0	4352	2	US-08-446-363-1	Sequence 1, Appli	681	13.2	66.0	8031	4	US-09-480-884A-254	Sequence 254, App
C 609	13.2	66.0	4352	5	US-08-316-950-16	Sequence 16, Appli	682	13.2	66.0	8366	4	US-09-542-615A-254	Sequence 254, App
C 610	13.2	66.0	4427	3	PCT-US95-12642-16	Sequence 16, Appli	683	13.2	66.0	8430	3	US-09-606-441B-254	Sequence 254, App
C 611	13.2	66.0	4505	1	US-09-342-353-2	Sequence 2, Appli	684	13.2	66.0	8430	3	US-09-872-733A-16	Sequence 6, Appli
					US-08-316-950-13	Sequence 13, Appli						US-09-131-028A-6	Sequence 10, Appli

c 665 13.2 66.0 8501 3 US-08-793-900-1  
c 666 13.2 66.0 8967 2 US-08-853-659A-6  
c 667 13.2 66.0 8967 2 US-08-853-659A-9  
c 668 13.2 66.0 8967 2 US-08-853-659A-64  
c 669 13.2 66.0 8967 2 US-08-853-659A-67  
c 670 13.2 66.0 9048 3 US-08-973-273-4  
c 671 13.2 66.0 10254 4 US-08-961-527-29  
c 672 13.2 66.0 10317 4 US-09-058-746-1  
c 673 13.2 66.0 10317 3 US-09-438-142-1  
c 674 13.2 66.0 10323 4 US-09-280-428A-11  
c 675 13.2 66.0 12284 4 US-08-876-991-1  
c 676 13.2 66.0 12284 4 US-09-059-853-1  
c 677 13.2 66.0 12311 3 US-08-750-717-1  
c 678 13.2 66.0 12655 4 US-08-961-527-134  
c 679 13.2 66.0 12846 4 US-09-553-690-1  
c 680 13.2 66.0 14113 3 US-09-223-134-1  
c 681 13.2 66.0 14113 3 US-08-992-801-1  
c 682 13.2 66.0 14113 3 US-09-223-535-1  
c 683 13.2 66.0 14194 4 US-09-574-424-3  
c 684 13.2 66.0 15538 4 US-09-554-337-1  
c 685 13.2 66.0 22846 2 US-08-469-461-3  
c 686 13.2 66.0 22846 3 US-07-890-609-3  
c 687 13.2 66.0 22846 3 US-08-846-762-1  
c 688 13.2 66.0 24701 2 US-08-853-659A-2  
c 689 13.2 66.0 24701 2 US-08-853-659A-3  
c 690 13.2 66.0 24701 2 US-08-853-659A-60  
c 691 13.2 66.0 24701 2 US-08-853-659A-61  
c 692 13.2 66.0 26016 4 US-09-326-480A-1  
c 693 13.2 66.0 28001 4 US-09-819-993-3  
c 694 13.2 66.0 28001 4 US-10-193-293-3  
c 695 13.2 66.0 28001 4 US-09-657-346A-96  
c 696 13.2 66.0 36138 4 US-08-311-731A-136  
c 697 13.2 66.0 40085 4 US-08-311-731A-26  
c 698 13.2 66.0 81001 4 US-09-750-580-1  
c 699 13.2 66.0 98844 4 US-09-791-211-10  
c 700 13.2 66.0 100848 4 US-09-596-002-39  
c 701 13.2 66.0 124884 4 US-09-661-596A-76  
c 702 13.2 66.0 124884 4 US-09-913-514-1  
c 703 13.2 66.0 125157 4 US-09-913-514-2  
c 704 13.2 66.0 125157 4 US-09-497-855A-38  
c 705 13.2 66.0 169989 4 US-09-676-610B-24  
c 706 13.2 66.0 197496 4 US-09-877-117A-10  
c 707 13.2 66.0 269223 4 US-09-596-002-41  
c 708 13.2 66.0 1664976 4 US-08-916-421B-1  
c 709 13.2 66.0 422 4 US-09-397-787-221  
c 710 13.2 66.0 474 4 US-09-621-976-19085  
c 711 13.2 66.0 702 4 US-09-328-352-3803  
c 712 13.2 66.0 1215 4 US-09-328-352-3884  
c 713 13.2 66.0 1257 2 US-08-546-117-8  
c 714 13.2 66.0 2196 4 US-09-620-312D-735  
c 715 13.2 66.0 2308 3 US-09-382-256-9  
c 716 13.2 66.0 2308 3 US-09-395-115-9  
c 717 13.2 66.0 2308 4 US-08-436-265-9  
c 718 13.2 66.0 2308 4 US-09-679-187-9  
c 719 13.2 66.0 2628 2 US-08-696-944-1  
c 720 13.2 66.0 2760 4 US-09-543-681A-86  
c 721 13.2 66.0 3713 1 US-08-330-537-3  
c 722 13.2 66.0 3713 1 US-08-330-537-4  
c 723 13.2 66.0 5474 1 US-08-137-252-2  
c 724 13.2 66.0 6396 4 US-09-620-312D-226  
c 725 13.2 66.0 18355 4 US-08-956-171E-67  
c 726 13.2 66.0 148567 4 US-10-254-869-3  
c 727 13.2 66.0 148567 4 US-09-801-876B-3  
c 728 13.2 66.0 21 4 US-09-422-978-9320  
c 729 13.2 66.0 201 4 US-09-404-879A-345  
c 730 13.2 66.0 231 4 US-09-252-991A-526  
c 731 13.2 66.0 258 4 US-09-543-681A-3278  
c 732 13.2 66.0 259 3 US-09-172-108-35  
c 733 13.2 66.0 259 3 US-09-172-108-35  
c 734 13.2 66.0 274 4 US-09-313-294A-1666  
c 735 13.2 66.0 279 4 US-09-489-039A-6634  
c 736 13.2 66.0 281 4 US-09-313-294A-1178  
c 737 13.2 66.0 298 4 US-09-313-294A-4029

Sequence 1, Appl1  
Sequence 6, Appl1  
Sequence 9, Appl1  
Sequence 64, Appl1  
Sequence 67, Appl1  
Sequence 4, Appl1  
Sequence 29, Appl1  
Sequence 1, Appl1  
Sequence 11, Appl1  
Sequence 1, Appl1  
Sequence 1, Appl1  
Sequence 134, App  
Sequence 1, Appl1  
Sequence 1, Appl1  
Sequence 1, Appl1  
Sequence 3, Appl1  
Sequence 3, Appl1  
Sequence 3, Appl1  
Sequence 1, Appl1  
Sequence 3, Appl1  
Sequence 2, Appl1  
Sequence 3, Appl1  
Sequence 60, Appl1  
Sequence 61, Appl1  
Sequence 1, Appl1  
Sequence 3, Appl1  
Sequence 3, Appl1  
Sequence 10, Appl1  
Sequence 39, Appl1  
Sequence 76, Appl1  
Sequence 1, Appl1  
Sequence 2, Appl1  
Sequence 38, Appl1  
Sequence 24, Appl1  
Sequence 10, Appl1  
Sequence 41, Appl1  
Sequence 1, Appl1  
Sequence 221, App  
Sequence 19085, A  
Sequence 3803, Ap  
Sequence 3884, Ap  
Sequence 8, Appl1  
Sequence 735, App  
Sequence 9, Appl1  
Sequence 9, Appl1  
Sequence 9, Appl1  
Sequence 9, Appl1  
Sequence 1, Appl1  
Sequence 86, Appl1  
Sequence 3, Appl1  
Sequence 4, Appl1  
Sequence 2, Appl1  
Sequence 26, App  
Sequence 67, App  
Sequence 3, Appl1  
Sequence 3, Appl1  
Sequence 9320, Ad  
Sequence 345, App  
Sequence 5226, Ap  
Sequence 3278, Ap  
Sequence 35, Appl1  
Sequence 33, Appl1  
Sequence 1666, Ap  
Sequence 6634, Ap  
Sequence 1178, Ap  
Sequence 4029, Ap

758 12.8 64.0 301 4 US-09-702-705-1714  
759 12.8 64.0 301 4 US-09-736-457-1714  
760 12.8 64.0 301 4 US-09-671-325-1714  
c 761 12.8 64.0 304 4 US-09-439-261-4  
c 762 12.8 64.0 304 4 US-09-227-613-4  
c 763 12.8 64.0 354 4 US-09-250-609-100  
c 764 12.8 64.0 354 4 US-09-250-611-100  
c 765 12.8 64.0 357 4 US-09-107-532A-93  
c 766 12.8 64.0 396 4 US-09-640-153-153  
c 767 12.8 64.0 396 4 US-09-713-550-153  
c 768 12.8 64.0 414 4 US-09-621-976-18437  
c 769 12.8 64.0 425 4 US-09-221-017B-344  
c 770 12.8 64.0 425 4 US-09-250-609-101  
c 771 12.8 64.0 425 4 US-09-250-611-101  
c 772 12.8 64.0 449 4 US-09-713-550-197  
c 773 12.8 64.0 454 3 US-08-991-789A-211  
c 774 12.8 64.0 454 4 US-09-062-451-211  
c 775 12.8 64.0 454 4 US-09-598-336-211  
c 776 12.8 64.0 454 4 US-09-289-198-211  
c 777 12.8 64.0 454 4 US-09-429-755-211  
c 778 12.8 64.0 477 4 US-09-252-991A-5351  
c 779 12.8 64.0 487 3 US-09-030-607-206  
c 780 12.8 64.0 487 4 US-09-439-313-206  
c 781 12.8 64.0 487 4 US-09-352-616A-206  
c 782 12.8 64.0 487 4 US-09-232-149A-206  
c 783 12.8 64.0 487 4 US-09-159-812-206  
c 784 12.8 64.0 487 4 US-09-636-215-206  
c 785 12.8 64.0 487 4 US-09-685-166A-206  
c 786 12.8 64.0 487 4 US-09-115-453-206  
c 787 12.8 64.0 487 4 US-09-688-489-206  
c 788 12.8 64.0 495 4 US-09-833-381-1146  
c 789 12.8 64.0 501 4 US-09-134-000C-2056  
c 790 12.8 64.0 506 4 US-09-645-168-7  
c 791 12.8 64.0 516 4 US-09-250-609-99  
c 792 12.8 64.0 516 4 US-09-250-611-99  
c 793 12.8 64.0 516 4 US-09-543-681A-2905  
c 794 12.8 64.0 527 4 US-09-621-976-1248  
c 795 12.8 64.0 540 4 US-09-134-000C-2406  
c 796 12.8 64.0 558 4 US-09-328-352-1847  
c 797 12.8 64.0 581 4 US-09-918-686-17  
c 798 12.8 64.0 586 4 US-09-465-559-21  
c 799 12.8 64.0 645 4 US-08-956-171E-686  
c 800 12.8 64.0 651 4 US-09-328-352-2818

ALIGNMENTS

US-09-651-656-8/c  
; Sequence 8, Application US/09651656  
; Patent No. 6340566  
; GENERAL INFORMATION:  
; APPLICANT: MCCUTCHEN-MALONEY, SANDRA  
; TITLE OF INVENTION: LIVERMORE NATIONAL LABORATORY  
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE  
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,  
; FILE REFERENCE: IT-10669  
; CURRENT APPLICATION NUMBER: US/09/651,656  
; CURRENT FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: 60/192,764  
; PRIOR FILING DATE: 2000-03-28  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 8  
; TYPE: DNA  
; LENGTH: 822  
; ORGANISM: Homo sapiens  
US-09-651-656-8

Query Match 100.0%; Score 20; DB 4; Length 822;  
Beet local Similarity 100.0%; Pred. No. 0.48;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGATG 20  
|||||  
Db 743 GGTCATCTCATGTGATG 724

RESULT 2  
US-09-650-855-8/c  
; Sequence 8, Application US/09650855  
; Patent No. 6365355  
; GENERAL INFORMATION:  
; APPLICANT: MCCUTHEN-MALONEY, SANDRA  
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY  
; TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTITATION OF DNA  
; TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA  
; FILE REFERENCE: IL-10284  
; CURRENT APPLICATION NUMBER: US/09/650,855  
; PRIOR FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: 60/192,764  
; PRIOR FILING DATE: 2000-03-28  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 822  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-650-855-8

Query Match 100.0%; Score 20; DB 4; Length 822;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGATG 20  
|||||  
Db 743 GGTCATCTCATGTGATG 724

RESULT 3  
US-08-961-527-40/c  
; Sequence 40, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunesh  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14273 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-40

Query Match 84.0%; Score 16.8; DB 4; Length 14273;  
Best Local Similarity 90.0%; Pred. No. 31;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGATG 20  
|||||  
Db 13385 GGTCATCTCATGTGATG 13366

RESULT 4  
US-09-385-982-62  
; Sequence 62, Application US/09385982  
; Patent No. 6262334  
; GENERAL INFORMATION:  
; APPLICANT: ENDEGE, WILSON O., ET AL.  
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
; TITLE OF INVENTION: PRODUCTS: II  
; FILE REFERENCE: CCNDA-260XX  
; CURRENT APPLICATION NUMBER: US/09/385,982  
; CURRENT FILING DATE: 1999-08-30  
; EARLIER APPLICATION NUMBER: 09/328,111  
; EARLIER FILING DATE: 1999-06-08  
; EARLIER APPLICATION NUMBER: 60/117,393  
; EARLIER FILING DATE: 1999-01-27  
; EARLIER APPLICATION NUMBER: 60/098,639  
; EARLIER FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 544  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 62  
; LENGTH: 614  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(614)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-385-982-62

Query Match 79.0%; Score 15.8; DB 3; Length 614;  
Best Local Similarity 89.5%; Pred. No. 60;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTGATG 20  
|||||  
Db 466 GTCCATCTCATGTGATG 484

RESULT 5  
US-09-023-655-1225/c  
; Sequence 1225, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Selhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9246741
; US-09-023-655-1225

```

```

Query Match          79.0%; Score 15.8; DB 4; Length 686;
Best Local Similarity 89.5%; Pred. No. 61;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 GGTCCATCTCATGTTGAT 19
        |||
Db      628 GGTCCATCTCATGTTGAT 610

```

```

RESULT 6
US-09-543-681A-137/C
; Sequence 137, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 137
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-137

```

```

Query Match          79.0%; Score 15.8; DB 4; Length 801;
Best Local Similarity 89.5%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2 GTTCATCTCATGTTGATG 20
        |||
Db      136 GACCAACTCATGTTGATG 118

```

```

RESULT 7
US-09-543-681A-1773/C
; Sequence 1773, Application US/09543681A
; Patent No. 6605709

```

```

; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1773
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-1773

```

```

Query Match          79.0%; Score 15.8; DB 4; Length 1176;
Best Local Similarity 89.5%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2 GTTCATCTCATGTTGATG 20
        |||
Db      417 GTTCATCTCATGTTGATG 399

```

```

RESULT 8
US-09-134-001C-388
; Sequence 388, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 388
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-388

```

```

Query Match          77.0%; Score 15.4; DB 4; Length 675;
Best Local Similarity 94.1%; Pred. No. 96;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 GTTCATCTCATGTTGA 18
        |||
Db      427 GTTCATCTCATGTTGA 443

```

```

RESULT 9
US-08-956-171E-1137
; Sequence 1137, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesh
; GIL H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville

```

STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 1137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 330 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1137:  
US-08-956-171E-1137

Query Match 76.0%; Score 15.2; DB 4; Length 330;  
Best Local Similarity 85.0%; Pred. No. 1.1e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTGATG 20  
DB 284 GGACCAACTCATGTGTGATG 303

RESULT 10  
US-09-489-039A-5224/C  
Sequence 5224 Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 5224  
LENGTH: 1419  
TYPE: DNA  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5224

Query Match 76.0%; Score 15.2; DB 4; Length 1419;  
Best Local Similarity 85.0%; Pred. No. 1.4e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTGATG 20  
DB 487 GGTTCATCTCATGTGTGCTG 468

RESULT 11  
US-09-328-352-2478

Sequence 2478 Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 2478  
LENGTH: 4884  
TYPE: DNA  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-2478

Query Match 76.0%; Score 15.2; DB 4; Length 4884;  
Best Local Similarity 85.0%; Pred. No. 1.7e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTGATG 20  
DB 1128 GTTCTTACTCATGTGTGATG 1147

RESULT 12  
US-08-646-695-1  
Sequence 1 Application US/08646695  
Patent No. 6168943  
GENERAL INFORMATION:  
APPLICANT: Rose, John K.  
TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,695  
FILING DATE: On Even Date Herewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mastrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 6523-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1431 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 760..2025  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2092..2886  
FEATURE:  
NAME/KEY: CDS

/ LOCATION: 2946..3632  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 3774..5306  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 5429..11755  
/ US-08-646-695-1

Query Match 76.0%; Score 15.2; DB 3; Length 14311;  
Best Local Similarity 85.0%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTTGAG 20  
DB 11703 GGTCATCTCATGTTGAG 11722

RESULT 13  
US-08-646-695-7/c

/ Sequence 7, Application US/08646695

/ Patent No. 6168943

/ GENERAL INFORMATION:

/ APPLICANT: ROSE, John K.

/ TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR

/ NUMBER OF SEQUENCES: 44

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: PENNIE & EDMONDS

/ STREET: 1155 Avenue of the Americas

/ CITY: New York

/ STATE: New York

/ COUNTRY: USA

/ ZIP: 10036-2711

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/646,695

/ FILING DATE: On Even Date Herewith

/ CLASSIFICATION: 435

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Mistrock, S. Leslie

/ REGISTRATION NUMBER: 18,872

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (212) 790-9090

/ TELEFAX: (212) 869-9741/8864

/ TELEX: 66141 PENNIE

/ INFORMATION FOR SEQ ID NO: 7:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 14311 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: double

/ TOPOLOGY: linear

/ MOLECULE TYPE: DNA

/ US-08-646-695-7

Query Match 76.0%; Score 15.2; DB 3; Length 14311;  
Best Local Similarity 85.0%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTTGAG 20  
DB 2609 GGTCATCTCATGTTGAG 2590

RESULT 14  
PCT-US96-06053-1

/ Sequence 1, Application PC/TUS9606053

/ GENERAL INFORMATION:

/ APPLICANT: Yale University  
/ TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR  
/ NUMBER OF SEQUENCES: 41  
/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: PENNIE & EDMONDS

/ STREET: 1155 Avenue of the Americas

/ CITY: New York

/ STATE: New York

/ COUNTRY: USA

/ ZIP: 10036-2711

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent Release #1.0, Version #1.30

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: PCT/US96/06053

/ FILING DATE: 01-MAY-1996

/ CLASSIFICATION:

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Mistrock, S. Leslie

/ REGISTRATION NUMBER: 18,872

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (212) 790-9090

/ TELEFAX: (212) 869-9741/8864

/ TELEX: 66141 PENNIE

/ INFORMATION FOR SEQ ID NO: 1:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 14311 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: double

/ TOPOLOGY: linear

/ MOLECULE TYPE: DNA

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: 760..2028

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: 2092..2889

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: 2946..3635

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: 3774..5309

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: 5429..11758

/ PCT-US96-06053-1

Query Match 76.0%; Score 15.2; DB 5; Length 14311;  
Best Local Similarity 85.0%; Pred. No. 2e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTTGAG 20  
DB 11703 GGTCATCTCATGTTGAG 11722

RESULT 15  
PCT-US96-06053-7/c

/ Sequence 7, Application PC/TUS9606053

/ GENERAL INFORMATION:

/ APPLICANT: Yale University

/ TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR

/ NUMBER OF SEQUENCES: 41

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: PENNIE & EDMONDS

/ STREET: 1155 Avenue of the Americas

/ CITY: New York

```
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06053
FILING DATE: 01-MAY-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-009-228
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 14311 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US96-06053-7

Query Match          76.0%; Score 15.2; DB 5; Length 14311;
Best Local Similarity 85.0%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GGTCCACTCTCATGTTGATG 20
Db      2609 GGTCTTACTCATGTTGAAG 2590

RESULT 16
US-08-311-731A-132
Sequence 132, Application US/08311731A
Patent No. 6583266
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 132:
```

```
SEQUENCE CHARACTERISTICS:
LENGTH: 36412 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-132

Query Match          76.0%; Score 15.2; DB 4; Length 36412;
Best Local Similarity 85.0%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GGTCCACTCTCATGTTGATG 20
Db      14170 GGACCACTACTCTCTGCATG 14189

RESULT 17
US-09-198-452A-1/c
Sequence 1, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(15000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (105001)..(120000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (120001)..(135000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (135001)..(150000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
LOCATION: (150001)..(165000)
OTHER INFORMATION: n=a or c or g or t
NAME/KEY: misc_feature
```



[illegible]

```
/ NAME/KEY: misc_feature
/ LOCATION: (900001) ..(915000)
/ OTHER INFORMATION: n=a or c or g or t
/ NAME/KEY: misc_feature
```

```
Query Match          76.0%; Score 15.2; DB 4; Length 1230025;
Best Local Similarity 85.0%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCTCATGTTGATG 20
      |||||
Db       175246 GGGCCGTTCTCATGTTGATG 175227
```

```
RESULT 18
US-09-103-840A-2/c
/ Sequence 2, Application US/09103840A
/ Patent No. 6294328
/ GENERAL INFORMATION:
/ APPLICANT: FLEISCHMAN, Robert D.
/ APPLICANT: WHITE, Owen R.
/ APPLICANT: FRASER, Claire M.
/ APPLICANT: VENTER, John C.
/ TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
/ FILE REFERENCE: 24366-20007.00
/ CURRENT APPLICATION NUMBER: US/09/103,840A
/ CURRENT FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 4403765
/ TYPE: DNA
/ ORGANISM: Mycobacterium tuberculosis
/ FEATURE:
/ OTHER INFORMATION: CDC 1551
/ OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2
```

```
Query Match          76.0%; Score 15.2; DB 3; Length 4403765;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCTCATGTTGATG 20
      |||||
Db       2185302 GGTCCAGACTCATGGGATG 2185283
```

```
RESULT 19
US-09-103-840A-1/c
/ Sequence 1, Application US/09103840A
/ Patent No. 6294328
/ GENERAL INFORMATION:
/ APPLICANT: FLEISCHMAN, Robert D.
/ APPLICANT: WHITE, Owen R.
/ APPLICANT: FRASER, Claire M.
/ APPLICANT: VENTER, John C.
/ TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
/ FILE REFERENCE: 24366-20007.00
/ CURRENT APPLICATION NUMBER: US/09/103,840A
/ CURRENT FILING DATE: 1998-06-24
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 4411529
/ TYPE: DNA
/ ORGANISM: Mycobacterium tuberculosis
/ OTHER INFORMATION: H37Rv
US-09-103-840A-1
```

```
Query Match          76.0%; Score 15.2; DB 3; Length 4411529;
```

```
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCTCATGTTGATG 20
      |||||
Db       2188003 GGTCCAGACTCATGGGATG 2187984
```

```
RESULT 20
US-09-328-352-1128
/ Sequence 1128, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: GARY L. BRETON et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 1128
/ LENGTH: 315
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-1128
```

```
Query Match          74.0%; Score 14.8; DB 4; Length 315;
Best Local Similarity 88.9%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      3 TCCATCTCATGTTGATG 20
      |||||
Db       263 TCTATGATCATGTTGATG 280
```

```
RESULT 21
US-09-328-352-1227/c
/ Sequence 1227, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: GARY L. BRETON et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 1227
/ LENGTH: 348
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-1227
```

```
Query Match          74.0%; Score 14.8; DB 4; Length 348;
Best Local Similarity 88.9%; Pred. No. 1.7e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      3 TCCATCTCATGTTGATG 20
      |||||
Db       305 TCTATGATCATGTTGATG 288
```

```
RESULT 22
US-09-621-976-3741
/ Sequence 3741, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: JOBERT, S.
/ APPLICANT: GIORDANO, J.Y.
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: GENSET.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621,976
```

```
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 3741
/ LENGTH: 387
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 191..376
US-09-621-976-3741

Query Match
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TCCATCATCATGTTGATG 20
Db 182 TCCAGATCATCATGATG 199

RESULT 23
US-09-328-352-992
/ Sequence 992, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 992
/ LENGTH: 546
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-992

Query Match
Best Local Similarity 88.9%; Pred. No. 1.9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 TCCATCATCATGTTGATG 20
Db 10 TCTATGATCATGTTGATG 27

RESULT 24
US-09-489-039A-911
/ Sequence 911, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 911
/ LENGTH: 2244
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-911

Query Match
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20
```

```
Db 1711 CATGCTCATGTTGATG 1726

RESULT 25
US-09-489-039A-953/C
/ Sequence 953, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 953
/ LENGTH: 2310
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-953

Query Match
Best Local Similarity 93.8%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20
Db 969 CATGCTCATGTTGATG 954

RESULT 26
US-08-136-743B-40
/ Sequence 40, Application US/08136743B
/ Patent No. 5459063
/ GENERAL INFORMATION:
/ APPLICANT: Barry S. Cooperman, Harvey Rubin,
/ APPLICANT: Jerome Salem, and Allison L. Fisher
/ TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
/ TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit
/ TITLE OF INVENTION: Thereof"
/ NUMBER OF SEQUENCES: 67
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: The University of Pennsylvania
/ STREET: Suite 330
/ STREET: 3700 Market Street
/ CITY: Philadelphia
/ STATE: Pennsylvania
/ COUNTRY: U.S.A.
/ ZIP: 19104-3246
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
/ COMPUTER: IBM PS/2
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/136,743B
/ FILING DATE: 10/14/93
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Monaco, Daniel A.
/ REGISTRATION NUMBER: 30,480
/ REFERENCE/DOCKET NUMBER: 3957-10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (215) 568-8383
/ TELEFAX: (215) 568-5549
/ TELEX: No. 5459063e
/ INFORMATION FOR SEQ ID NO: 40:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2433 nucleotides
/ TYPE: nucleic acid
```

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-136-743B-40

Query Match 72.0%; Score 14.4; DB 1; Length 2433;  
Best Local Similarity 93.8%; Pred. No. 3.8e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATCTGTG 17  
DB 2272 GTCCATCTCATCTGTG 2287

RESULT 27  
US-08-136-743B-3  
Sequence 3, Application US/08136743B  
Patent No. 5459063  
GENERAL INFORMATION:  
APPLICANT: Barry S. Cooperman, Harvey Rubin,  
Jerome Salem, and Allison L. Fisher  
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-  
TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: The University of Pennsylvania  
STREET: Suite 330  
STREET: 3700 Market Street  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19104-3246  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/136,743B  
FILING DATE: 10/14/93  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 3957-10  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5459063e  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2663 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-136-743B-3

Query Match 72.0%; Score 14.4; DB 1; Length 2663;  
Best Local Similarity 93.8%; Pred. No. 3.8e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATCTGTG 17  
DB 2499 GTCCATCTCATCTGTG 2514

RESULT 28  
US-08-247-901C-1  
Sequence 1, Application US/08247901C  
Patent No. 5750384  
GENERAL INFORMATION:  
APPLICANT: Jacobs et al  
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS

NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ameter, Rothstein & Ebenstein  
STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,901C  
FILING DATE: May 23, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/057,531  
FILING DATE: April 29, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Elizabeth A  
REGISTRATION NUMBER: 39,911  
REFERENCE/DOCKET NUMBER: 96700/273  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50341  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: L5 shuttle phasmid sequence  
HYPOTHETICAL: NO  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: L5 mycobacteriophage  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION: No. 5750384e  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-247-901C-1

Query Match 72.0%; Score 14.4; DB 1; Length 50341;  
Best Local Similarity 93.8%; Pred. No. 6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



ORGANISM: mycobacteriophage L5  
STRAIN: not applicable  
INDIVIDUAL ISOLATE: L5  
DEVELOPMENTAL STAGE: not applicable  
HAPLOTYPE: not applicable  
TISSUE TYPE: not applicable  
CELL LINE: not applicable  
ORGANELLE: not applicable  
IMMEDIATE SOURCE: mycobacteriophage L5 particles  
POSITION IN GENOME: entire genome  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS: Hatfull and Sarkis  
TITLE: DNA Sequence, Structure and Gene  
TITLE: Expression of Mycobacteriophage L5:  
TITLE: A Phage System for Mycobacterial  
TITLE: Genetics  
JOURNAL: Molecular Microbiology  
VOLUME: 7  
PAGES: 395-405  
DATE: 1993  
US-09-426-436-1

Query Match 72.0%; Score 14.4; DB 3; Length 52297;  
Best Local Similarity 93.8%; Pred. No. 6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20  
DB 9132 CATCTCATGTTGATG 9147

RESULT 31  
US-08-705-557-1  
Sequence 1, Application US/08705557  
Patent No. 6300061  
GENERAL INFORMATION:  
APPLICANT: William R. Jacobs, Jr.  
APPLICANT: Barry R. Bloom  
APPLICANT: Graham F. Hatfull  
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC  
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ametier, Rothstein & Ebenstein  
STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/705,557  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/057,531  
FILING DATE:  
APPLICATION NUMBER: 07/833,431  
FILING DATE: February 7, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pasqualini, Patricia A.  
REGISTRATION NUMBER: 34,894  
REFERENCE/DOCKET NUMBER: 96700/238

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 52297  
TYPE: nucleotide  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: phage genome sequence  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: not applicable.  
ORIGINAL SOURCE:  
ORGANISM: mycobacteriophage L5  
STRAIN: not applicable  
INDIVIDUAL ISOLATE: L5  
DEVELOPMENTAL STAGE: not applicable  
HAPLOTYPE: not applicable  
TISSUE TYPE: not applicable  
CELL TYPE: not applicable  
CELL LINE: not applicable  
ORGANELLE: not applicable  
IMMEDIATE SOURCE: mycobacteriophage L5 particles  
POSITION IN GENOME: entire genome  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS: Hatfull and Sarkis  
TITLE: DNA Sequence, Structure and Gene  
TITLE: Expression of Mycobacteriophage L5:  
TITLE: A Phage System for Mycobacterial  
TITLE: Genetics  
JOURNAL: Molecular Microbiology  
VOLUME: 7  
PAGES: 395-405  
DATE: 1993  
US-08-705-557-1

Query Match 72.0%; Score 14.4; DB 4; Length 52297;  
Best Local Similarity 93.8%; Pred. No. 6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CATACTCATGTTGATG 20  
DB 9132 CATCTCATGTTGATG 9147

RESULT 32  
US-09-135-782-10  
Sequence 10, Application US/09135782  
Patent No. 6027929  
GENERAL INFORMATION:  
APPLICANT: Xu, Shuang-Yong  
TITLE OF INVENTION: Method For Cloning And Producing The NspI Restriction  
TITLE OF INVENTION: Endonuclease in E. coli and Purification Of The  
TITLE OF INVENTION: Recombinant NspI Restriction Endonuclease  
FILE REFERENCE: NEB-143  
CURRENT APPLICATION NUMBER: US/09/135,782  
CURRENT FILING DATE: 1998-08-18  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 36  
TYPE: DNA  
ORGANISM: No. 6027929loc sp.  
US-09-135-782-10





SEQ ID NO 13151  
LENGTH: 483  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-13151

Query Match 71.0%; Score 14.2; DB 4; Length 483;  
Best Local Similarity 84.2%; Pred. No. 3.7e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGAT 19  
Db 26 GATCCATGTCATGTTGAT 44

RESULT 37  
US-09-134-000C-244  
Sequence 244, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 244  
LENGTH: 510  
TYPE: DNA  
ORGANISM: Enterococcus faecalis  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (330)..(491)  
OTHER INFORMATION: Nucleotides 330, 382, 430, 440, 449 & 491 are "n" wherein  
OTHER INFORMATION: "n" = any nucleotide.  
OTHER INFORMATION: 1  
US-09-134-000C-244

Query Match 71.0%; Score 14.2; DB 4; Length 510;  
Best Local Similarity 84.2%; Pred. No. 3.7e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GTCCTACTCATGTTGAT 20  
Db 187 GTCGATCTCATGTTGAT 205

RESULT 38  
US-09-107-532A-735  
Sequence 735, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 735:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1..555  
SEQUENCE DESCRIPTION: SEQ ID NO: 735:  
US-09-107-532A-735

Query Match 71.0%; Score 14.2; DB 4; Length 555;  
Best Local Similarity 84.2%; Pred. No. 3.8e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGAT 19  
Db 189 GTCATCTCGATGTTGAT 207

RESULT 39  
US-09-342-681C-11/c  
Sequence 11, Application US/09342681C  
Patent No. 6355782  
GENERAL INFORMATION:  
APPLICANT: Zonana et al.  
TITLE OF INVENTION: Hypodermic ectodermal dysplasia genes and proteins  
FILE REFERENCE: 52978  
CURRENT APPLICATION NUMBER: US/09/342,681C  
CURRENT FILING DATE: 1999-06-29  
PRIOR APPLICATION NUMBER: 60/092,279  
PRIOR FILING DATE: 1998-07-09  
PRIOR APPLICATION NUMBER: 60/112,366  
PRIOR FILING DATE: 1998-12-15  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 743  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (302)  
OTHER INFORMATION: n represents a, c, t, or g  
US-09-342-681C-11

Query Match 71.0%; Score 14.2; DB 4; Length 743;  
Best Local Similarity 84.2%; Pred. No. 3.9e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GTCCTACTCATGTTGAT 20  
Db 538 GTCCTGCTCATGTTGAT 520

RESULT 40  
US-09-252-991A-9153/C  
; Sequence 9153, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 9153  
; LENGTH: 909  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-9153

Query Match 71.0%; Score 14.2; DB 4; Length 909;  
Best Local Similarity 84.2%; Pred. No. 4.1e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GTCCATCTCATGTTGATG 20  
DB 265 GTCCATGCCCATGCTGATG 247

RESULT 41  
US-09-634-238-121  
; Sequence 121, Application US/09634238  
; Patent No. 6544772  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Havukkala, Ilkka J.  
; APPLICANT: Bloksberg, Leonard, N.  
; APPLICANT: Lubbers, Mark W.  
; APPLICANT: Dekker, James  
; APPLICANT: Christenson, Anna C.  
; APPLICANT: Holland, Ross  
; APPLICANT: O'Toole, Paul W.  
; APPLICANT: Reid, Julian R.  
; APPLICANT: Coolbear, Timothy  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; FILE REFERENCE: 11000.1043U1  
; CURRENT APPLICATION NUMBER: US/09/634,238  
; PRIOR FILING DATE: 2000-08-08  
; NUMBER OF SEQ ID NOS: 422  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 121  
; LENGTH: 1065  
; TYPE: DNA  
; ORGANISM: Lactobacillus rhamnosus  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(1065)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-634-238-121

Query Match 71.0%; Score 14.2; DB 4; Length 1065;  
Best Local Similarity 84.2%; Pred. No. 4.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GGTCATCTCATGTTGAT 19  
DB 224 GGTCATCTCATGCGGAT 242

RESULT 42  
US-09-342-681C-14/C  
; Sequence 14, Application US/09342681C  
; Patent No. 6355782  
; GENERAL INFORMATION:  
; APPLICANT: Zonana et al.  
; TITLE OF INVENTION: Hypohydratic ectodermal dysplasia genes and proteins  
; FILE REFERENCE: 52978  
; CURRENT APPLICATION NUMBER: US/09/342,681C  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 60/092,279  
; PRIOR FILING DATE: 1998-07-09  
; PRIOR APPLICATION NUMBER: 60/112,366  
; PRIOR FILING DATE: 1998-12-15  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1176)  
US-09-342-681C-14

Query Match 71.0%; Score 14.2; DB 4; Length 1176;  
Best Local Similarity 84.2%; Pred. No. 4.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GTCCATCTCATGTTGATG 20  
DB 1128 GTGCTTGCTCATGTTGATG 1110

RESULT 43  
US-09-252-991A-11577/C  
; Sequence 11577, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 11577  
; LENGTH: 1185  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-11577

Query Match 71.0%; Score 14.2; DB 4; Length 1185;  
Best Local Similarity 84.2%; Pred. No. 4.2e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GTCCATCTCATGTTGATG 20  
DB 490 GTCCATGCTCATGTTGATG 472

RESULT 44  
US-08-286-020-1  
; Sequence 1, Application US/08286020  
; Patent No. 5539095  
; GENERAL INFORMATION:  
; APPLICANT: Masomeh B. Sticklen and Ravindra K. Hajela  
; TITLE OF INVENTION: A Chitinase cDNA Clone From a

```

; TITLE OF INVENTION: Disease Resistant American
; TITLE OF INVENTION: Elm tree
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: Acer
; OPERATING SYSTEM: MS-DOS 5.00
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,020
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: No. 55390956
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1225
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Ulmus Americana
; STRAIN: NPS 3-487
; INDIVIDUAL ISOLATE: N/A
; DEVELOPMENTAL STAGE: N/A
; HAPLOTYPE: N/A
; TISSUE TYPE: N/A
; CELL TYPE: N/A
; CELL LINE: N/A
; ORGANELLE: N/A
; IMMEDIATE SOURCE: N/A
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: chitinase encoding DNA
; LOCATION:
; IDENTIFICATION METHOD: sequencing
; OTHER INFORMATION: DNA needed for chitinase
; OTHER INFORMATION: in elm.
; PUBLICATION INFORMATION:
;
; US-08-286-020-1
;
; Query Match 71.0%; Score 14.2; DB 1; Length 1225;
; Best Local Similarity 84.2%; Pred. No. 4.3e+02;
; Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
; QY 1 GGTCATACATCTGTGAT 19
; |||||
; 517 GGTCATACATCTGTGGGAT 535
;
; RESULT 45
; US-08-603-919-1
; Sequence 1, Application US/08603919
```

```

; Patent No. 5728382
; GENERAL INFORMATION:
; APPLICANT: Masomeh B. Sticklen and Ravindra K. Hajela
; TITLE OF INVENTION: A Chitinase cDNA Clone From a
; TITLE OF INVENTION: Disease Resistant American
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48864
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: Acer
; OPERATING SYSTEM: MS-DOS 5.00
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/603,919
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ian C. McLeod
; REGISTRATION NUMBER: 20,931
; REFERENCE/DOCKET NUMBER: MSU 4.1-207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 347-4100
; TELEFAX: (517) 347-4103
; TELEX: No. 5728382e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1225
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Ulmus Americana
; STRAIN: NPS 3-487
; INDIVIDUAL ISOLATE: N/A
; DEVELOPMENTAL STAGE: N/A
; HAPLOTYPE: N/A
; TISSUE TYPE: N/A
; CELL TYPE: N/A
; CELL LINE: N/A
; ORGANELLE: N/A
; IMMEDIATE SOURCE: N/A
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: chitinase encoding DNA
; LOCATION:
; IDENTIFICATION METHOD: sequencing
; OTHER INFORMATION: DNA needed for chitinase
; OTHER INFORMATION: in elm.
; PUBLICATION INFORMATION:
;
; US-08-603-919-1
;
; Query Match 71.0%; Score 14.2; DB 1; Length 1225;
; Best Local Similarity 84.2%; Pred. No. 4.3e+02;
; Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
; QY 1 GGTCATACATCTGTGAT 19
; |||||
; 517 GGTCATACATCTGTGGGAT 535
```

```

RESULT 46
US-08-343-101A-8/c
: Sequence 8, Application US/08343101A
: Patent No. 5830759
: GENERAL INFORMATION:
: APPLICANT: Chang, Yuan
: APPLICANT: Moore, Patrick S.
: TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
: TITLE OF INVENTION: Virus Sequences And Uses Thereof
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/343.101A
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: White Eqd., John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 45185-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-278-0400
: TELEFAX: 212-391-0526
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1315 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: N
: ANTI-SENSE: N
: US-08-343-101A-8

Query Match 71.0%; Score 14.2; DB 2; Length 1315;
Best Local Similarity 84.2%; Pred.No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0

QY 2 GTCCATACTCATGTGTGATG 20
      ||| ||| ||| ||| |||
Db 805 GTCCCACTGATGTGTGATG 787

RESULT 47
US-09-183-688-8/c
: Sequence 8, Application US/09183688
: Patent No. 6093550
: GENERAL INFORMATION:
: APPLICANT: Chang, Yuan
: APPLICANT: Moore, Patrick S.
: TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma
: TITLE OF INVENTION: Virus Sequences And Uses Thereof
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/183,688
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/343,101
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: White Esq., John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 45185-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-278-0400
: TELEFAX: 212-391-0526
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1315 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: N
: ANTI-SENSE: N
:
: US-09-183-688-8
:
: Query Match 71.0%; Score 14.2; DB 3; Length 1315;
: Best Local Similarity 84.2%; Pred. No. 4.3e+02;
: Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: QY 2 GTCCATCTCATGTTGATG 20
: ||||| ||||| ||||| |||||
: Db 805 GTCCCACTGATGTTGATG 787
:
: RESULT 48
: US-09-519-489-8/C
: Sequence 8, Application US/09519489
: Patent No. 6500663
: GENERAL INFORMATION:
: APPLICANT: Chang, Yuan
: APPLICANT: Moore, Patrick S.
: TITLE OF INVENTION: Unique Sequences Associated Kaposi's Sarcoma
: TITLE OF INVENTION: Virus Sequences And Uses Thereof
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/519,489
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/343,101
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: White Esq., John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 45185-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-278-0400
: TELEFAX: 212-391-0526
:
:

```

/ INFORMATION FOR SEQ ID NO: 8:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 1315 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: DNA (genomic)  
/ HYPOTHETICAL: N  
/ ANTI-SENSE: N  
US-09-519-489-8

Query Match 71.0%; Score 14.2; DB 4; Length 1315;  
Best Local Similarity 84.2%; Pred. No. 4.3e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTGTGATG 20  
DB 805 GTCCCACTCATGTGTGATG 787

RESULT 49  
US-09-252-991A-11863  
/ Sequence 11863, Application US/09252991A  
/ Patent No. 6551795  
/ GENERAL INFORMATION:  
/ APPLICANT: Marc J. Rubenfield et al.  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
/ FILE REFERENCE: 107196.136  
/ CURRENT APPLICATION NUMBER: US/09/252,991A  
/ CURRENT FILING DATE: 1999-02-18  
/ PRIOR APPLICATION NUMBER: US 60/074,788  
/ PRIOR FILING DATE: 1998-02-18  
/ PRIOR APPLICATION NUMBER: US 60/094,190  
/ PRIOR FILING DATE: 1998-07-27  
/ NUMBER OF SEQ ID NOS: 33142  
/ SEQ ID NO 11863  
/ LENGTH: 1386  
/ TYPE: DNA  
/ ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-11863

Query Match 71.0%; Score 14.2; DB 4; Length 1386;  
Best Local Similarity 84.2%; Pred. No. 4.3e+02;  
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTGTGATG 20  
DB 1248 GTCCATCTCATGTGTGATG 1266

RESULT 50  
US-09-489-039A-3739  
/ Sequence 3739, Application US/09489039A  
/ Patent No. 6610836  
/ GENERAL INFORMATION:  
/ APPLICANT: Gary Breton et. al  
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
/ FILE REFERENCE: 2709.2004001  
/ CURRENT APPLICATION NUMBER: US/09/489,039A  
/ CURRENT FILING DATE: 2000-01-27  
/ PRIOR APPLICATION NUMBER: US 60/117,747  
/ PRIOR FILING DATE: 1999-01-29  
/ NUMBER OF SEQ ID NOS: 14342  
/ SEQ ID NO 3739  
/ LENGTH: 1398  
/ TYPE: DNA  
/ ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-3739

Query Match 71.0%; Score 14.2; DB 4; Length 1398;  
Best Local Similarity 84.2%; Pred. No. 4.4e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GTCCATCTCATGTGTGATG 20  
DB 433 GTCCCACTCATGTGTGATG 451

Search completed: August 17, 2004, 15:46:45  
Job time : 126 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 15:02:58 ; Search time 362 Seconds  
(without alignments)  
271.083 Million cell updates/sec

Title: US-09-825-489-3  
Perfect score: 20  
Sequence: 1 ggtccatccatcgttgatg 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 800 summaries

Database :

Published Applications NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	11	US-09-825-489-3
2	20	100.0	486	9	US-09-867-701-7544
3	20	100.0	506	9	US-09-867-701-2117
4	20	100.0	513	9	US-09-867-701-2107
5	20	100.0	1377	11	US-09-825-489-13
6	20	100.0	1407	15	US-10-103-313-116
7	20	100.0	1470	15	US-10-103-313-116
8	17	85.0	1526	13	US-10-027-632-257254
9	17	85.0	1526	16	US-10-027-632-257254
10	16.8	84.0	14273	13	US-10-158-844-40
11	16.4	82.0	1049	13	US-10-027-632-253417
12	16.4	82.0	1049	16	US-10-027-632-253417
13	15.8	79.0	238	13	US-10-085-783A-40734
14	15.8	79.0	238	16	US-10-242-535A-40734

C 15	15.8	79.0	296	12	US-09-732-627A-2949	Sequence 2949, App
C 16	15.8	79.0	431	15	US-10-060-036-3118	Sequence 3118, App
C 17	15.8	79.0	462	10	US-09-918-995-12216	Sequence 12216, A
C 18	15.8	79.0	540	13	US-10-085-783A-45577	Sequence 45577, A
C 19	15.8	79.0	540	16	US-10-242-535A-45577	Sequence 45577, A
C 20	15.8	79.0	614	10	US-09-871-161-62	Sequence 62, Appl
C 21	15.8	79.0	686	17	US-10-641-643-1225	Sequence 1225, App
C 22	15.8	79.0	765	13	US-10-424-599-20394	Sequence 20394, A
C 23	15.8	79.0	789	17	US-10-437-663-64475	Sequence 64475, A
C 24	15.8	79.0	1329	17	US-10-437-663-7531	Sequence 7531, App
C 25	15.8	79.0	3021	12	US-10-041-018-111	Sequence 111, App
C 26	15.8	79.0	5874	12	US-10-041-018-112	Sequence 112, App
C 27	15.8	79.0	101209	13	US-10-087-192-460	Sequence 460, App
C 28	15.4	77.0	519	13	US-10-027-632-79848	Sequence 79848, A
C 29	15.4	77.0	519	13	US-10-027-632-79848	Sequence 79849, A
C 30	15.4	77.0	519	16	US-10-027-632-79848	Sequence 79848, A
C 31	15.4	77.0	519	16	US-10-027-632-79849	Sequence 79849, A
C 32	15.4	77.0	551	9	US-09-764-847-497	Sequence 497, App
C 33	15.4	77.0	551	15	US-10-092-154-497	Sequence 497, App
C 34	15.4	77.0	573	13	US-10-027-632-66707	Sequence 66707, A
C 35	15.4	77.0	573	13	US-10-027-632-66707	Sequence 66708, A
C 36	15.4	77.0	573	13	US-10-027-632-311196	Sequence 311196, A
C 37	15.4	77.0	573	13	US-10-027-632-311197	Sequence 311197, A
C 38	15.4	77.0	573	16	US-10-027-632-66707	Sequence 66707, A
C 39	15.4	77.0	573	16	US-10-027-632-66707	Sequence 66708, A
C 40	15.4	77.0	573	16	US-10-027-632-311196	Sequence 311196, A
C 41	15.4	77.0	573	16	US-10-027-632-311197	Sequence 311197, A
C 42	15.4	77.0	715	13	US-10-424-599-103050	Sequence 103050, A
C 43	15.4	77.0	3767	13	US-10-424-599-23361	Sequence 23361, A
C 44	15.4	77.0	119414	13	US-10-087-192-1807	Sequence 1807, App
C 45	15.4	77.0	335913	10	US-09-754-853A-2	Sequence 2, Appl
C 46	15.4	77.0	335913	10	US-09-754-853A-3	Sequence 3, Appl
C 47	15.2	76.0	292	13	US-10-424-599-127062	Sequence 127062, A
C 48	15.2	76.0	330	8	US-08-781-986A-1137	Sequence 1137, App
C 49	15.2	76.0	330	13	US-10-329-624-1137	Sequence 1137, App
C 50	15.2	76.0	358	13	US-10-424-599-130364	Sequence 130364, A
C 51	15.2	76.0	449	17	US-10-437-663-31302	Sequence 31302, A
C 52	15.2	76.0	464	13	US-10-424-599-90875	Sequence 90875, A
C 53	15.2	76.0	469	13	US-10-424-599-32789	Sequence 32789, A
C 54	15.2	76.0	572	13	US-10-027-632-208409	Sequence 208409, A
C 55	15.2	76.0	572	16	US-10-027-632-208409	Sequence 208409, A
C 56	15.2	76.0	606	17	US-10-437-663-55289	Sequence 55289, A
C 57	15.2	76.0	630	13	US-10-424-599-128976	Sequence 128976, A
C 58	15.2	76.0	1035	16	US-10-312-773-218	Sequence 218, App
C 59	15.2	76.0	1107	13	US-10-282-122A-26487	Sequence 26487, A
C 60	15.2	76.0	1110	13	US-10-424-599-79603	Sequence 79603, A
C 61	15.2	76.0	1287	9	US-09-738-626-2128	Sequence 2128, App
C 62	15.2	76.0	1335	13	US-10-627-476-621	Sequence 621, App
C 63	15.2	76.0	1410	13	US-10-627-476-619	Sequence 619, App
C 64	15.2	76.0	1513	13	US-10-027-632-253949	Sequence 253949, A
C 65	15.2	76.0	1513	16	US-10-027-632-253949	Sequence 253949, A
C 66	15.2	76.0	1908	15	US-10-156-761-5276	Sequence 5276, App
C 67	15.2	76.0	2000	11	US-09-938-842A-4942	Sequence 4942, App
C 68	15.2	76.0	2000	9	US-09-938-842A-4942	Sequence 4942, App
C 69	15.2	76.0	2913	17	US-10-437-663-17890	Sequence 17890, A
C 70	15.2	76.0	2916	12	US-10-052-482-117	Sequence 117, App
C 71	15.2	76.0	2916	17	US-10-394-948-27	Sequence 27, Appl
C 72	15.2	76.0	3334	10	US-09-822-946-102	Sequence 102, App
C 73	15.2	76.0	3670	15	US-10-205-823-125	Sequence 125, App
C 74	15.2	76.0	3717	17	US-10-437-663-91561	Sequence 91561, A
C 75	15.2	76.0	3892	12	US-10-052-482-116	Sequence 116, App
C 76	15.2	76.0	3892	17	US-10-394-948-26	Sequence 26, Appl
C 77	15.2	76.0	6537	12	US-09-997-722-24	Sequence 24, Appl
C 78	15.2	76.0	6537	15	US-10-175-523-131	Sequence 131, App
C 79	15.2	76.0	6537	15	US-10-175-523-131	Sequence 186, App
C 80	15.2	76.0	7254	12	US-10-267-502-151	Sequence 151, App
C 81	15.2	76.0	17177	15	US-10-225-810-7	Sequence 7, Appl
C 82	15.2	76.0	36605	13	US-10-087-192-304	Sequence 304, App
C 83	15.2	76.0	82993	15	US-10-080-170-645	Sequence 645, App
C 84	15.2	76.0	82993	17	US-10-080-170-645	Sequence 645, App
C 85	15.2	76.0	87311	13	US-10-087-192-1417	Sequence 1417, App
C 86	15.2	76.0	96598	12	US-10-052-482-115	Sequence 115, App
C 87	15.2	76.0	96598	17	US-10-394-948-25	Sequence 25, Appl



C 88	15.2	76.0	326014	9	US-09-731-231A-3	Sequence 3, Appl1
C 89	15.2	76.0	326014	17	US-10-751-985-3	Sequence 3, Appl1
C 90	15.2	76.0	1230025	16	US-10-289-762-1	Sequence 1, Appl1
C 91	15.2	76.0	3309400	9	US-09-738-626-1	Sequence 1, Appl1
C 92	15.2	76.0	9025608	15	US-10-156-761-1	Sequence 961, App
C 93	15	75.0	331	13	US-10-424-559-961	Sequence 7781, Ap
C 94	15	75.0	593	9	US-09-864-761-7781	Sequence 10, Appl
C 95	15	75.0	1851	10	US-09-898-200-10	Sequence 2, Appl1
C 96	15	75.0	2000	13	US-09-898-200-2	Sequence 2025, Ap
C 97	14.8	74.0	138	13	US-10-282-122A-2025	Sequence 15648, A
C 98	14.8	74.0	380	9	US-09-864-761-15648	Sequence 32154, A
C 99	14.8	74.0	380	9	US-09-864-761-32154	Sequence 6873, Ap
C 100	14.8	74.0	414	13	US-10-282-122A-6873	Sequence 1021, App
C 101	14.8	74.0	459	9	US-09-864-761-102	Sequence 133049, A
C 102	14.8	74.0	463	13	US-10-424-559-133049	Sequence 2262, Ap
C 103	14.8	74.0	499	13	US-10-343-887-2262	Sequence 2262, Ap
C 104	14.8	74.0	499	13	US-10-172-118-2262	Sequence 37952, A
C 105	14.8	74.0	503	13	US-10-027-632-37952	Sequence 37953, A
C 106	14.8	74.0	503	13	US-10-027-632-37953	Sequence 65145, A
C 107	14.8	74.0	503	13	US-10-027-632-65145	Sequence 65146, A
C 108	14.8	74.0	503	13	US-10-027-632-65146	Sequence 310829, A
C 109	14.8	74.0	503	13	US-10-027-632-310829	Sequence 310830, A
C 110	14.8	74.0	503	13	US-10-027-632-310830	Sequence 37952, A
C 111	14.8	74.0	503	16	US-10-027-632-37952	Sequence 37953, A
C 112	14.8	74.0	503	16	US-10-027-632-37953	Sequence 65145, A
C 113	14.8	74.0	503	16	US-10-027-632-65145	Sequence 65146, A
C 114	14.8	74.0	503	16	US-10-027-632-65146	Sequence 310829, A
C 115	14.8	74.0	503	16	US-10-027-632-310829	Sequence 310830, A
C 116	14.8	74.0	503	16	US-10-027-632-310830	Sequence 2076, Ap
C 117	14.8	74.0	504	17	US-10-021-332-2076	Sequence 271233, A
C 118	14.8	74.0	515	13	US-10-027-632-271233	Sequence 50695, A
C 119	14.8	74.0	515	16	US-10-027-632-271233	Sequence 50695, A
C 120	14.8	74.0	528	13	US-10-027-632-50695	Sequence 88554, A
C 121	14.8	74.0	528	16	US-10-027-632-50695	Sequence 88554, A
C 122	14.8	74.0	536	13	US-10-027-632-88554	Sequence 88554, A
C 123	14.8	74.0	536	16	US-10-027-632-88554	Sequence 68799, A
C 124	14.8	74.0	536	17	US-10-437-963-68799	Sequence 4237, Ap
C 125	14.8	74.0	564	15	US-10-029-386-4237	Sequence 20547, A
C 126	14.8	74.0	577	9	US-09-864-761-20547	Sequence 12543, A
C 127	14.8	74.0	594	13	US-10-425-114-12543	Sequence 75460, A
C 128	14.8	74.0	596	13	US-10-027-632-75460	Sequence 75460, A
C 129	14.8	74.0	596	16	US-10-027-632-75460	Sequence 208952, A
C 130	14.8	74.0	600	13	US-10-027-632-208951	Sequence 208952, A
C 131	14.8	74.0	600	13	US-10-027-632-208952	Sequence 208952, A
C 132	14.8	74.0	600	16	US-10-027-632-208951	Sequence 208952, A
C 133	14.8	74.0	600	16	US-10-027-632-208952	Sequence 246341, A
C 134	14.8	74.0	610	13	US-10-027-632-246341	Sequence 246341, A
C 135	14.8	74.0	610	13	US-10-027-632-246342	Sequence 246342, A
C 136	14.8	74.0	610	16	US-10-027-632-246342	Sequence 86952, A
C 137	14.8	74.0	610	16	US-10-027-632-86952	Sequence 86952, A
C 138	14.8	74.0	626	13	US-10-027-632-86952	Sequence 86952, A
C 139	14.8	74.0	626	16	US-10-027-632-86952	Sequence 89394, A
C 140	14.8	74.0	630	13	US-10-027-632-89394	Sequence 89394, A
C 141	14.8	74.0	630	13	US-10-027-632-89394	Sequence 89395, A
C 142	14.8	74.0	630	13	US-10-027-632-303989	Sequence 303989, A
C 143	14.8	74.0	630	13	US-10-027-632-303989	Sequence 303989, A
C 144	14.8	74.0	630	16	US-10-027-632-303990	Sequence 89394, A
C 145	14.8	74.0	630	16	US-10-027-632-89395	Sequence 89395, A
C 146	14.8	74.0	630	16	US-10-027-632-303989	Sequence 303989, A
C 147	14.8	74.0	630	16	US-10-027-632-303990	Sequence 303990, A
C 148	14.8	74.0	636	13	US-10-027-632-313264	Sequence 313264, A
C 149	14.8	74.0	636	16	US-10-027-632-313264	Sequence 19342, A
C 150	14.8	74.0	642	13	US-10-282-122A-19342	Sequence 279251, A
C 151	14.8	74.0	656	13	US-10-027-632-279251	Sequence 279251, A
C 152	14.8	74.0	656	16	US-10-027-632-279251	Sequence 211220, A
C 153	14.8	74.0	663	13	US-10-027-632-211220	Sequence 211220, A
C 154	14.8	74.0	663	16	US-10-027-632-3142283	Sequence 1442283, A
C 155	14.8	74.0	728	13	US-10-027-632-1442283	Sequence 1442283, A
C 156	14.8	74.0	728	13	US-10-027-632-1442284	Sequence 1442283, A
C 157	14.8	74.0	728	16	US-10-027-632-1442283	Sequence 1442283, A
C 158	14.8	74.0	728	16	US-10-027-632-1442284	Sequence 1442284, A
C 159	14.8	74.0	826	13	US-10-027-632-157166	Sequence 157166, A
C 160	14.8	74.0	826	13	US-10-027-632-157167	Sequence 157167, A
C 161	14.8	74.0	826	16	US-10-027-632-157166	Sequence 157166, A
C 162	14.8	74.0	826	16	US-10-027-632-157167	Sequence 8029, Ap
C 163	14.8	74.0	835	13	US-10-027-632-8029	Sequence 8029, Ap
C 164	14.8	74.0	835	16	US-10-027-632-8029	Sequence 529, App
C 165	14.8	74.0	858	13	US-10-343-550A-529	Sequence 533, App
C 166	14.8	74.0	858	13	US-10-343-550A-533	Sequence 54628, A
C 167	14.8	74.0	868	13	US-10-424-559-54628	Sequence 374, App
C 168	14.8	74.0	957	9	US-09-886-055-576	Sequence 376, App
C 169	14.8	74.0	957	9	US-09-886-055-576	Sequence 376, App
C 170	14.8	74.0	957	10	US-09-804-291-376	Sequence 100467, A
C 171	14.8	74.0	957	10	US-10-437-963-100467	Sequence 100866, A
C 172	14.8	74.0	1060	17	US-10-437-963-100866	Sequence 12041, A
C 173	14.8	74.0	1075	17	US-10-437-963-100866	Sequence 597, App
C 174	14.8	74.0	1181	13	US-10-425-114-12041	Sequence 599, App
C 175	14.8	74.0	1242	15	US-10-017-161-597	Sequence 214985, A
C 176	14.8	74.0	1258	15	US-10-017-161-597	Sequence 214985, A
C 177	14.8	74.0	1258	15	US-10-017-161-597	Sequence 523, App
C 178	14.8	74.0	1308	13	US-10-027-632-214985	Sequence 525, App
C 179	14.8	74.0	1308	16	US-10-027-632-214985	Sequence 135, App
C 180	14.8	74.0	1357	16	US-10-292-798-525	Sequence 42, Appl
C 181	14.8	74.0	1357	16	US-10-292-798-525	Sequence 42, Appl
C 182	14.8	74.0	1364	9	US-09-823-880A-135	Sequence 42, Appl
C 183	14.8	74.0	1568	9	US-09-925-299-42	Sequence 42, Appl
C 184	14.8	74.0	1568	10	US-09-925-299-42	Sequence 3781, App
C 185	14.8	74.0	1654	13	US-10-425-114-31792	Sequence 1804, Ap
C 186	14.8	74.0	1719	13	US-10-282-122A-40474	Sequence 2354, Ap
C 187	14.8	74.0	1818	13	US-10-282-122A-18936	Sequence 18936, A
C 188	14.8	74.0	1836	9	US-09-864-761-16939	Sequence 16939, A
C 189	14.8	74.0	1938	17	US-10-437-963-40328	Sequence 40328, A
C 190	14.8	74.0	1978	17	US-10-437-963-40328	Sequence 46591, A
C 191	14.8	74.0	1987	9	US-09-864-761-3781	Sequence 3781, App
C 192	14.8	74.0	2000	16	US-10-260-338-1804	Sequence 1804, Ap
C 193	14.8	74.0	2000	16	US-10-260-338-2354	Sequence 2354, Ap
C 194	14.8	74.0	2002	9	US-09-887-576-298	Sequence 298, App
C 195	14.8	74.0	2004	9	US-09-887-576-299	Sequence 299, App
C 196	14.8	74.0	2033	17	US-10-466-720-24	Sequence 24, Appl
C 197	14.8	74.0	2129	13	US-10-424-559-54625	Sequence 54625, A
C 198	14.8	74.0	2379	11	US-09-938-842A-2171	Sequence 2171, Ap
C 199	14.8	74.0	2379	11	US-09-938-842A-2171	Sequence 2171, Ap
C 200	14.8	74.0	2850	13	US-10-027-632-111951	Sequence 111951, A
C 201	14.8	74.0	2850	16	US-10-027-632-111951	Sequence 6150, App
C 202	14.8	74.0	3061	13	US-10-425-114-6150	Sequence 5504, A
C 203	14.8	74.0	3134	13	US-10-424-559-45533	Sequence 52305, A
C 204	14.8	74.0	3256	17	US-10-437-963-25305	Sequence 26879, A
C 205	14.8	74.0	3273	16	US-10-369-493-26879	Sequence 125, App
C 206	14.8	74.0	3295	16	US-10-330-797-125	Sequence 961, App
C 207	14.8	74.0	3545	15	US-10-101-464A-951	Sequence 61646, A
C 208	14.8	74.0	3765	17	US-10-437-963-61646	Sequence 26395, A
C 209	14.8	74.0	3909	17	US-10-437-963-26395	Sequence 45318, A
C 210	14.8	74.0	4564	17	US-10-437-963-45318	Sequence 1531, Ap
C 211	14.8	74.0	4595	9	US-09-917-800A-1531	Sequence 9854, Ap
C 212	14.8	74.0	5085	15	US-10-198-464-9854	Sequence 9854, Ap
C 213	14.8	74.0	5110	15	US-10-101-464A-963	Sequence 963, App
C 214	14.8	74.0	5119	17	US-10-437-963-80960	Sequence 80960, A
C 215	14.8	74.0	6106	16	US-10-062-674-1648	Sequence 1648, Ap
C 216	14.8	74.0	7566	10	US-09-764-991-10010	Sequence 10010, A
C 217	14.8	74.0	7741	13	US-10-092-000A-75	Sequence 75, Appl
C 218	14.8	74.0	7983	15	US-10-017-161-1987	Sequence 1987, Ap
C 219	14.8	74.0	29283	16	US-10-252-798-1635	Sequence 1635, Ap
C 220	14.8	74.0	32757	17	US-10-332-881-140	Sequence 140, App
C 221	14.8	74.0	33675	9	US-09-921-992-2	Sequence 2, Appl1
C 222	14.8	74.0	35216	9	US-09-747-810-1	Sequence 1, Appl1
C 223	14.8	74.0	90468	16	US-10-085-117-160	Sequence 160, App
C 224	14.8	74.0	94917	13	US-10-087-192-184	Sequence 184, App
C 225	14.8	74.0	105519	13	US-10-087-192-658	Sequence 658, App
C 226	14.8	74.0	126512	9	US-09-804-474A-3	Sequence 3, Appl1
C 227	14.8	74.0	135827	17	US-10-332-881-232	Sequence 232, App
C 228	14.8	74.0	137049	13	US-10-087-192-11198	Sequence 1198, App
C 229	14.8	74.0	173808	14	US-10-003-806-10	Sequence 10, Appl
C 230	14.8	74.0	178996	17	US-10-450-826-53	Sequence 53, Appl
C 231	14.8	74.0	296405	13	US-10-087-192-1036	Sequence 1036, App
C 232	14.8	74.0	344548	13	US-10-087-192-334	Sequence 334, App
C 233	14.8	74.0	418550	16	US-10-292-798-1463	Sequence 1463, App

C 234	14.8	74.0	684973	9	US-09-263-959-1	Sequence 1, Appl1	C 307	14.2	71.0	171	15	US-10-029-386-27336	Sequence 27336, A
C 235	14.8	74.0	1691139	15	US-10-067-514-1	Sequence 1, Appl1	C 308	14.2	71.0	174	13	US-10-085-783A-35599	Sequence 35599, A
C 236	14.8	74.0	1691139	16	US-10-449-723-1	Sequence 1, Appl1	C 309	14.2	71.0	174	16	US-10-242-535A-35599	Sequence 35599, A
C 237	14.4	72.0	220	13	US-10-424-599-44369	Sequence 44369, A	C 310	14.2	71.0	176	13	US-10-085-783A-17286	Sequence 17286, A
C 238	14.4	72.0	275	17	US-10-437-963-10372	Sequence 10372, A	C 311	14.2	71.0	176	16	US-10-242-535A-17286	Sequence 17286, A
C 239	14.4	72.0	353	13	US-10-424-599-99713	Sequence 99713, A	C 312	14.2	71.0	190	13	US-10-085-783A-22589	Sequence 22589, A
C 240	14.4	72.0	392	10	US-09-918-995-4410	Sequence 4410, Ap	C 313	14.2	71.0	190	16	US-10-242-535A-22589	Sequence 22589, A
C 241	14.4	72.0	487	13	US-10-424-599-10890	Sequence 10890, A	C 314	14.2	71.0	209	9	US-09-960-352-8245	Sequence 8245, Ap
C 242	14.4	72.0	576	9	US-09-777-564-121	Sequence 121, App	C 315	14.2	71.0	218	15	US-10-106-698-1739	Sequence 1739, Ap
C 243	14.4	72.0	576	15	US-10-015-219-121	Sequence 121, App	C 316	14.2	71.0	225	17	US-10-437-663-16904	Sequence 16904, A
C 244	14.4	72.0	582	9	US-09-777-564-555	Sequence 555, App	C 317	14.2	71.0	227	13	US-10-085-783A-9450	Sequence 9450, Ap
C 245	14.4	72.0	582	15	US-10-015-219-555	Sequence 555, App	C 318	14.2	71.0	227	16	US-10-242-535A-9450	Sequence 9450, Ap
C 246	14.4	72.0	594	13	US-10-027-632-96747	Sequence 96747, A	C 319	14.2	71.0	234	15	US-10-029-986-14659	Sequence 14659, A
C 247	14.4	72.0	594	16	US-10-027-632-306306	Sequence 306306, A	C 320	14.2	71.0	234	16	US-10-393-840-809	Sequence 809, App
C 248	14.4	72.0	594	13	US-10-027-632-96747	Sequence 96747, A	C 321	14.2	71.0	247	9	US-09-864-761-20089	Sequence 20089, A
C 249	14.4	72.0	594	16	US-10-027-632-306306	Sequence 306306, A	C 322	14.2	71.0	249	13	US-10-424-599-96366	Sequence 46366, A
C 250	14.4	72.0	604	17	US-10-021-323-15825	Sequence 15825, A	C 323	14.2	71.0	255	10	US-09-930-213-768	Sequence 768, App
C 251	14.4	72.0	608	13	US-10-027-632-208171	Sequence 208171, A	C 324	14.2	71.0	273	13	US-10-085-783A-13656	Sequence 13656, A
C 252	14.4	72.0	608	13	US-10-027-632-208172	Sequence 208172, A	C 325	14.2	71.0	273	16	US-10-242-535A-13656	Sequence 13656, A
C 253	14.4	72.0	608	13	US-10-027-632-208173	Sequence 208173, A	C 326	14.2	71.0	298	9	US-09-960-352-7665	Sequence 7665, Ap
C 254	14.4	72.0	608	13	US-10-027-632-208174	Sequence 208174, A	C 327	14.2	71.0	298	13	US-10-424-599-140679	Sequence 140679, A
C 255	14.4	72.0	608	13	US-10-027-632-208175	Sequence 208175, A	C 328	14.2	71.0	305	13	US-10-424-599-4051	Sequence 4051, Ap
C 256	14.4	72.0	608	13	US-10-027-632-208176	Sequence 208176, A	C 329	14.2	71.0	315	12	US-09-922-393-1129	Sequence 1129, Ap
C 257	14.4	72.0	608	16	US-10-027-632-208171	Sequence 208171, A	C 330	14.2	71.0	329	8	US-08-781-986A-986	Sequence 986, App
C 258	14.4	72.0	608	16	US-10-027-632-208172	Sequence 208172, A	C 331	14.2	71.0	329	13	US-10-329-624-886	Sequence 986, App
C 259	14.4	72.0	608	16	US-10-027-632-208173	Sequence 208173, A	C 332	14.2	71.0	330	9	US-09-764-869-445	Sequence 445, App
C 260	14.4	72.0	608	16	US-10-027-632-208174	Sequence 208174, A	C 333	14.2	71.0	330	9	US-09-764-869-182	Sequence 182, App
C 261	14.4	72.0	608	16	US-10-027-632-208175	Sequence 208175, A	C 334	14.2	71.0	330	10	US-09-989-442-71	Sequence 71, Appl1
C 262	14.4	72.0	608	16	US-10-027-632-208176	Sequence 208176, A	C 335	14.2	71.0	330	15	US-10-091-504-445	Sequence 445, App
C 263	14.4	72.0	636	13	US-10-027-632-254270	Sequence 254270, A	C 336	14.2	71.0	330	16	US-10-227-577-445	Sequence 445, App
C 264	14.4	72.0	636	16	US-10-027-632-254270	Sequence 254270, A	C 337	14.2	71.0	333	13	US-10-424-599-124672	Sequence 124672, A
C 265	14.4	72.0	654	13	US-10-027-632-217620	Sequence 217620, A	C 338	14.2	71.0	339	15	US-10-029-986-32457	Sequence 32457, A
C 266	14.4	72.0	654	13	US-10-027-632-217621	Sequence 217621, A	C 339	14.2	71.0	341	13	US-10-424-599-11581	Sequence 11581, A
C 267	14.4	72.0	654	16	US-10-027-632-217622	Sequence 217622, A	C 340	14.2	71.0	346	13	US-10-424-599-138636	Sequence 138636, A
C 268	14.4	72.0	654	16	US-10-027-632-217620	Sequence 217620, A	C 341	14.2	71.0	360	13	US-10-424-599-9746	Sequence 9746, Ap
C 269	14.4	72.0	654	16	US-10-027-632-217621	Sequence 217621, A	C 342	14.2	71.0	360	13	US-10-085-783A-48729	Sequence 48729, A
C 270	14.4	72.0	654	16	US-10-027-632-217622	Sequence 217622, A	C 343	14.2	71.0	360	16	US-10-242-535A-48729	Sequence 48729, A
C 271	14.4	72.0	659	17	US-10-021-333-6838	Sequence 6838, Ap	C 344	14.2	71.0	363	13	US-10-424-599-73201	Sequence 73201, A
C 272	14.4	72.0	1155	9	US-09-810-936-329	Sequence 329, App	C 345	14.2	71.0	370	13	US-10-085-783A-110	Sequence 110, App
C 273	14.4	72.0	1155	9	US-09-924-400-329	Sequence 329, App	C 346	14.2	71.0	370	16	US-10-242-535A-110	Sequence 110, App
C 274	14.4	72.0	1155	12	US-10-079-117B-329	Sequence 329, App	C 347	14.2	71.0	373	13	US-10-085-783A-57310	Sequence 57310, A
C 275	14.4	72.0	1155	15	US-10-212-679-329	Sequence 329, App	C 348	14.2	71.0	373	16	US-10-242-535A-57310	Sequence 57310, A
C 276	14.4	72.0	1155	17	US-10-451-168-37	Sequence 37, Appl1	C 349	14.2	71.0	376	13	US-10-424-599-5095	Sequence 5095, Ap
C 277	14.4	72.0	1226	13	US-10-027-632-252975	Sequence 252975, A	C 350	14.2	71.0	376	13	US-10-085-783A-25270	Sequence 25270, A
C 278	14.4	72.0	1226	16	US-10-027-632-252975	Sequence 252975, A	C 351	14.2	71.0	376	16	US-10-242-535A-25270	Sequence 25270, A
C 279	14.4	72.0	1233	13	US-10-425-114-20605	Sequence 20605, A	C 352	14.2	71.0	381	13	US-10-282-122A-36660	Sequence 36660, A
C 280	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 353	14.2	71.0	382	9	US-09-960-352-11499	Sequence 11499, A
C 281	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 354	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 282	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 355	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 283	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 356	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 284	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 357	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 285	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 358	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 286	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 359	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 287	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 360	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 288	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 361	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 289	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 362	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 290	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 363	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 291	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 364	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 292	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 365	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 293	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 366	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 294	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 367	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 295	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 368	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 296	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 369	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 297	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 370	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 298	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 371	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 299	14.4	72.0	1240	13	US-10-424-599-91781	Sequence 91781, A	C 372	14.2	71.0	384	13	US-10-282-122A-38815	Sequence 38815, A
C 300	14.2	71.0	20	16	US-10-199-199-54	Sequence 54, Appl1	C 373	14.2	71.0	442	16	US-10-242-535A-56594	Sequence 56594, A
C 301	14.2	71.0	65	10	US-10-199-199-123	Sequence 123, App	C 374	14.2	71.0	450	17	US-10-437-963-33014	Sequence 33014, A
C 302	14.2	71.0	130	15	US-10-029-386-17205	Sequence 17205, A	C 375	14.2	71.0	457	17	US-10-437-963-78936	Sequence 78936, A
C 303	14.2	71.0	160	9	US-09-864-761-31676	Sequence 31676, A	C 376	14.2	71.0	460	13	US-10-085-783A-25991	Sequence 25991, A
C 304	14.2	71.0	165	9	US-09-864-761-23502	Sequence 23502, A	C 377	14.2	71.0	466	16	US-10-242-535A-25991	Sequence 25991, A
C 305	14.2	71.0	168	15	US-10-029-386-15999	Sequence 15999, A	C 378	14.2	71.0	468	13	US-10-085-783A-26172	Sequence 26172, A
C 306	14.2	71.0	170	9	US-09-989-920-151	Sequence 151, App	C 379	14.2	71.0	468	16	US-10-242-535A-26172	Sequence 26172, A

C 380	14.2	71.0	470	13	US-10-085-783A-42081	Sequence 42081, A	C 453	14.2	71.0	828	16	US-10-027-632-173547	Sequence 173547, A
C 381	14.2	71.0	470	16	US-10-248-535A-42081	Sequence 42081, A	C 454	14.2	71.0	828	16	US-10-027-632-173548	Sequence 173548, A
C 382	14.2	71.0	471	9	US-09-783-590-10451	Sequence 10451, A	C 455	14.2	71.0	828	16	US-10-027-632-173549	Sequence 173549, A
C 383	14.2	71.0	473	9	US-09-864-761-15149	Sequence 15149, A	C 456	14.2	71.0	842	13	US-10-425-114-21555	Sequence 21555, A
C 384	14.2	71.0	476	13	US-10-027-632-181465	Sequence 181465, A	C 457	14.2	71.0	870	13	US-10-282-122A-27655	Sequence 27655, A
C 385	14.2	71.0	476	13	US-10-027-632-181466	Sequence 181466, A	C 458	14.2	71.0	884	16	US-10-369-993-30485	Sequence 30485, A
C 386	14.2	71.0	476	16	US-10-027-632-181465	Sequence 181465, A	C 459	14.2	71.0	912	13	US-10-424-599-26124	Sequence 26124, A
C 387	14.2	71.0	476	16	US-10-027-632-181466	Sequence 181466, A	C 460	14.2	71.0	915	13	US-10-282-122A-39868	Sequence 39868, A
C 388	14.2	71.0	477	10	US-09-918-995-12069	Sequence 12069, A	C 461	14.2	71.0	928	13	US-10-425-114-19003	Sequence 19003, A
C 389	14.2	71.0	482	10	US-09-918-995-13549	Sequence 13549, A	C 462	14.2	71.0	928	13	US-10-282-122A-27931	Sequence 27931, A
C 390	14.2	71.0	484	9	US-09-864-761-3314	Sequence 3314, A	C 463	14.2	71.0	934	13	US-10-425-114-20130	Sequence 20130, A
C 391	14.2	71.0	486	9	US-09-864-761-5456	Sequence 5456, A	C 464	14.2	71.0	936	13	US-10-425-114-3330	Sequence 3330, A
C 392	14.2	71.0	506	9	US-09-796-692-2838	Sequence 2838, A	C 465	14.2	71.0	936	13	US-10-425-114-14387	Sequence 14187, A
C 393	14.2	71.0	506	15	US-10-040-862-2838	Sequence 2838, A	C 466	14.2	71.0	954	13	US-10-425-114-29727	Sequence 29727, A
C 394	14.2	71.0	506	16	US-10-057-475B-2838	Sequence 2838, A	C 467	14.2	71.0	970	13	US-10-342-887-2667	Sequence 2667, A
C 395	14.2	71.0	506	16	US-10-155-884B-2838	Sequence 2838, A	C 468	14.2	71.0	970	13	US-10-172-118-2667	Sequence 2667, A
C 396	14.2	71.0	507	15	US-10-028-386-9845	Sequence 9845, A	C 469	14.2	71.0	993	15	US-10-156-761-4178	Sequence 4709, A
C 397	14.2	71.0	521	13	US-10-027-632-312912	Sequence 312912, A	C 470	14.2	71.0	1008	17	US-10-437-563-26282	Sequence 26292, A
C 398	14.2	71.0	521	16	US-10-027-632-312912	Sequence 312912, A	C 471	14.2	71.0	1018	13	US-10-027-632-10345	Sequence 10345, A
C 399	14.2	71.0	534	15	US-10-029-386-13636	Sequence 13636, A	C 472	14.2	71.0	1018	13	US-10-027-632-10345	Sequence 10345, A
C 400	14.2	71.0	548	9	US-09-864-761-6769	Sequence 6769, A	C 473	14.2	71.0	1018	16	US-10-027-632-10346	Sequence 10346, A
C 401	14.2	71.0	548	13	US-10-425-114-16250	Sequence 16250, A	C 474	14.2	71.0	1018	16	US-10-027-632-10346	Sequence 10346, A
C 402	14.2	71.0	548	15	US-10-028-386-850	Sequence 850, A	C 475	14.2	71.0	1025	9	US-09-971-536-25	Sequence 25, A
C 403	14.2	71.0	549	15	US-10-029-386-956	Sequence 956, A	C 476	14.2	71.0	1045	13	US-10-425-114-17568	Sequence 17568, A
C 404	14.2	71.0	554	17	US-10-437-963-73593	Sequence 73593, A	C 477	14.2	71.0	1059	13	US-10-425-114-11667	Sequence 11667, A
C 405	14.2	71.0	558	15	US-10-029-386-2299	Sequence 2299, A	C 478	14.2	71.0	1098	13	US-10-425-114-5886	Sequence 5886, A
C 406	14.2	71.0	560	15	US-10-029-386-3505	Sequence 3505, A	C 479	14.2	71.0	1104	13	US-10-424-599-35651	Sequence 35651, A
C 407	14.2	71.0	567	13	US-10-240-425-439	Sequence 439, A	C 480	14.2	71.0	1106	13	US-10-027-632-117171	Sequence 11711, A
C 408	14.2	71.0	568	13	US-10-027-632-49222	Sequence 49222, A	C 481	14.2	71.0	1106	13	US-10-027-632-117171	Sequence 11712, A
C 409	14.2	71.0	568	13	US-10-027-632-63505	Sequence 63505, A	C 482	14.2	71.0	1106	16	US-10-027-632-117171	Sequence 11711, A
C 410	14.2	71.0	568	13	US-10-027-632-88517	Sequence 88517, A	C 483	14.2	71.0	1106	16	US-10-027-632-117172	Sequence 11712, A
C 411	14.2	71.0	568	16	US-10-027-632-49222	Sequence 49222, A	C 484	14.2	71.0	1146	13	US-10-282-122A-20688	Sequence 20688, A
C 412	14.2	71.0	568	16	US-10-027-632-63505	Sequence 63505, A	C 485	14.2	71.0	1148	13	US-10-225-060A-991	Sequence 991, A
C 413	14.2	71.0	568	16	US-10-027-632-88517	Sequence 88517, A	C 486	14.2	71.0	1148	16	US-10-374-780A-2813	Sequence 2813, A
C 414	14.2	71.0	589	13	US-10-027-632-24849	Sequence 24849, A	C 487	14.2	71.0	1156	13	US-10-027-632-210835	Sequence 210835, A
C 415	14.2	71.0	589	13	US-10-027-632-248500	Sequence 248500, A	C 488	14.2	71.0	1156	16	US-10-027-632-210835	Sequence 210835, A
C 416	14.2	71.0	589	13	US-10-027-632-248501	Sequence 248501, A	C 489	14.2	71.0	1176	10	US-09-729-658B-14	Sequence 14, A
C 417	14.2	71.0	589	16	US-10-027-632-248499	Sequence 248499, A	C 490	14.2	71.0	1187	13	US-10-424-599-50893	Sequence 50893, A
C 418	14.2	71.0	589	16	US-10-027-632-248500	Sequence 248500, A	C 491	14.2	71.0	1206	16	US-10-369-493-32245	Sequence 32245, A
C 419	14.2	71.0	589	16	US-10-027-632-248501	Sequence 248501, A	C 492	14.2	71.0	1214	9	US-09-764-846-52	Sequence 52, A
C 420	14.2	71.0	625	17	US-10-437-963-63535	Sequence 63535, A	C 493	14.2	71.0	1214	15	US-10-091-483-52	Sequence 52, A
C 421	14.2	71.0	625	13	US-10-027-632-17865	Sequence 17865, A	C 494	14.2	71.0	1248	13	US-10-282-122A-36127	Sequence 36127, A
C 422	14.2	71.0	622	13	US-10-027-632-178694	Sequence 178694, A	C 495	14.2	71.0	1254	17	US-10-437-963-43994	Sequence 43994, A
C 423	14.2	71.0	622	16	US-10-027-632-178685	Sequence 178685, A	C 496	14.2	71.0	1271	13	US-10-027-632-121656	Sequence 212656, A
C 424	14.2	71.0	622	16	US-10-027-632-178694	Sequence 178694, A	C 497	14.2	71.0	1271	16	US-10-027-632-121656	Sequence 212656, A
C 425	14.2	71.0	627	13	US-10-027-632-5709	Sequence 5709, A	C 498	14.2	71.0	1281	13	US-10-282-122A-27814	Sequence 27814, A
C 426	14.2	71.0	627	16	US-10-027-632-5709	Sequence 5709, A	C 499	14.2	71.0	1287	13	US-10-282-122A-18884	Sequence 18884, A
C 427	14.2	71.0	637	13	US-10-027-632-190606	Sequence 190606, A	C 500	14.2	71.0	1290	5	US-09-764-846-124	Sequence 124, A
C 428	14.2	71.0	637	13	US-10-027-632-230816	Sequence 230816, A	C 501	14.2	71.0	1290	15	US-10-091-483-124	Sequence 124, A
C 429	14.2	71.0	637	16	US-10-027-632-190606	Sequence 190606, A	C 502	14.2	71.0	1314	16	US-10-369-493-24743	Sequence 24743, A
C 430	14.2	71.0	637	16	US-10-027-632-230816	Sequence 230816, A	C 503	14.2	71.0	1368	13	US-10-424-599-121015	Sequence 121015, A
C 431	14.2	71.0	643	13	US-10-027-632-238962	Sequence 238962, A	C 504	14.2	71.0	1407	9	US-10-425-114-20301	Sequence 20301, A
C 432	14.2	71.0	643	13	US-10-027-632-238962	Sequence 238962, A	C 505	14.2	71.0	1452	9	US-09-938-842A-2402	Sequence 2402, A
C 433	14.2	71.0	643	13	US-10-027-632-238963	Sequence 238963, A	C 506	14.2	71.0	1452	11	US-09-938-842A-2402	Sequence 2402, A
C 434	14.2	71.0	643	16	US-10-027-632-238961	Sequence 238961, A	C 507	14.2	71.0	1467	17	US-10-437-963-33825	Sequence 33825, A
C 435	14.2	71.0	643	16	US-10-027-632-238962	Sequence 238962, A	C 508	14.2	71.0	1470	9	US-09-938-842A-1796	Sequence 1796, A
C 436	14.2	71.0	644	16	US-10-027-632-238963	Sequence 238963, A	C 509	14.2	71.0	1470	11	US-09-938-842A-1796	Sequence 1796, A
C 437	14.2	71.0	644	13	US-10-027-632-94592	Sequence 94592, A	C 510	14.2	71.0	1574	10	US-09-729-658B-1	Sequence 1, A
C 438	14.2	71.0	644	13	US-10-027-632-305513	Sequence 305513, A	C 511	14.2	71.0	1591	16	US-10-320-997-1159	Sequence 1159, A
C 439	14.2	71.0	644	16	US-10-027-632-94592	Sequence 94592, A	C 512	14.2	71.0	1604	13	US-10-425-114-28650	Sequence 28650, A
C 440	14.2	71.0	644	16	US-10-027-632-305513	Sequence 305513, A	C 513	14.2	71.0	1637	13	US-10-425-114-20378	Sequence 20378, A
C 441	14.2	71.0	685	13	US-10-027-632-113381	Sequence 113381, A	C 514	14.2	71.0	1647	13	US-10-424-599-50848	Sequence 50848, A
C 442	14.2	71.0	685	16	US-10-027-632-113381	Sequence 113381, A	C 515	14.2	71.0	1651	10	US-09-533-029-59	Sequence 59, A
C 443	14.2	71.0	703	17	US-10-437-963-17503	Sequence 17503, A	C 516	14.2	71.0	1651	13	US-10-412-699B-881	Sequence 881, A
C 444	14.2	71.0	725	16	US-10-027-632-20217	Sequence 20217, A	C 517	14.2	71.0	1651	13	US-10-225-066A-299	Sequence 299, A
C 445	14.2	71.0	743	10	US-09-729-658B-11	Sequence 11, A	C 518	14.2	71.0	1651	16	US-10-374-780A-271	Sequence 271, A
C 446	14.2	71.0	752	9	US-09-822-849A-556	Sequence 556, A	C 519	14.2	71.0	1654	13	US-10-282-122A-27624	Sequence 27624, A
C 447	14.2	71.0	752	16	US-10-027-632-150015	Sequence 150015, A	C 520	14.2	71.0	1654	16	US-10-264-413-16	Sequence 16, A
C 448	14.2	71.0	826	13	US-10-027-632-150015	Sequence 150015, A	C 521	14.2	71.0	1725	13	US-10-282-122A-30934	Sequence 30934, A
C 449	14.2	71.0	826	13	US-10-027-632-13547	Sequence 13547, A	C 522	14.2	71.0	1745	13	US-10-424-599-38057	Sequence 38057, A
C 450	14.2	71.0	828	13	US-10-027-632-173547	Sequence 173547, A	C 523	14.2	71.0	1745	13	US-10-425-114-5089	Sequence 5089, A
C 451	14.2	71.0	828	13	US-10-027-632-173548	Sequence 173548, A	C 524	14.2	71.0	1747	13	US-10-425-114-19556	Sequence 19556, A
C 452	14.2	71.0	828	13	US-10-027-632-173549	Sequence 173549, A	C 525	14.2	71.0	1790	10	US-09-919-039-150	Sequence 150, A

C 526	14.2	71.0	1790	14	US-10-044-090-505	Sequence 505, App	C 599	14.2	71.0	3169	9	US-09-981-353-53	Sequence 53, App1
C 527	14.2	71.0	1838	13	US-10-425-114-5743	Sequence 573, Ap	C 600	14.2	71.0	3169	15	US-10-235-994-15	Sequence 15, App1
C 528	14.2	71.0	1842	16	US-10-429-873A-1	Sequence 1, App1	C 601	14.2	71.0	3181	14	US-10-025-167-18	Sequence 18, App1
C 529	14.2	71.0	1861	13	US-10-425-114-620	Sequence 620, App	C 602	14.2	71.0	3188	16	US-10-108-260A-581	Sequence 581, App
C 530	14.2	71.0	1879	17	US-10-647-426-30	Sequence 30, App1	C 603	14.2	71.0	3195	10	US-09-867-034-22	Sequence 22, App1
C 531	14.2	71.0	1883	15	US-10-210-120-98	Sequence 98, App1	C 604	14.2	71.0	3195	13	US-10-276-115-22	Sequence 22, App1
C 532	14.2	71.0	1890	16	US-10-159-563-390	Sequence 390, App	C 605	14.2	71.0	3199	15	US-10-158-646-39	Sequence 39, App1
C 533	14.2	71.0	1899	13	US-10-283-122A-9684	Sequence 9684, Ap	C 606	14.2	71.0	3196	13	US-10-276-774-93	Sequence 993, App1
C 534	14.2	71.0	1902	16	US-10-398-221-1054	Sequence 1054, Ap	C 607	14.2	71.0	3204	15	US-10-345-680-31	Sequence 31, App1
C 535	14.2	71.0	1902	16	US-10-398-221-2831	Sequence 2831, Ap	C 608	14.2	71.0	3207	15	US-10-101-510-660	Sequence 660, App1
C 536	14.2	71.0	1905	13	US-10-425-114-32032	Sequence 32032, A	C 609	14.2	71.0	3218	16	US-10-087-580-33	Sequence 33, App1
C 537	14.2	71.0	1908	13	US-10-282-122A-14840	Sequence 14840, A	C 610	14.2	71.0	3234	13	US-10-423-543-48	Sequence 48, App1
C 538	14.2	71.0	1954	17	US-10-641-643-286	Sequence 286, App	C 611	14.2	71.0	3262	13	US-10-369-022-57	Sequence 57, App1
C 539	14.2	71.0	1962	15	US-10-233-665-5	Sequence 5, App1	C 612	14.2	71.0	3262	16	US-10-199-199-4	Sequence 4, App1
C 540	14.2	71.0	1989	10	US-09-738-630-93	Sequence 93, App1	C 613	14.2	71.0	3265	9	US-09-989-732-378	Sequence 378, App
C 541	14.2	71.0	1989	15	US-10-175-523-190	Sequence 190, App	C 614	14.2	71.0	3265	9	US-09-989-723-378	Sequence 378, App
C 542	14.2	71.0	1989	15	US-10-102-524-1725	Sequence 1725, Ap	C 615	14.2	71.0	3265	9	US-09-988-279-378	Sequence 378, App
C 543	14.2	71.0	1989	15	US-10-232-665-7	Sequence 7, App1	C 616	14.2	71.0	3265	9	US-09-989-723-378	Sequence 378, App
C 544	14.2	71.0	2000	13	US-10-614-076-99	Sequence 99, App1	C 617	14.2	71.0	3265	9	US-09-989-731-378	Sequence 378, App
C 545	14.2	71.0	2000	16	US-10-260-238-1960	Sequence 1960, Ap	C 618	14.2	71.0	3265	9	US-09-989-732-378	Sequence 378, App
C 546	14.2	71.0	2002	9	US-09-935-390A-5	Sequence 5, App1	C 619	14.2	71.0	3265	9	US-09-991-073-378	Sequence 378, App
C 547	14.2	71.0	2016	16	US-10-369-493-37033	Sequence 37033, A	C 620	14.2	71.0	3265	9	US-09-990-442-378	Sequence 378, App
C 548	14.2	71.0	2050	13	US-10-614-076-101	Sequence 101, App	C 621	14.2	71.0	3265	9	US-09-991-163-378	Sequence 378, App
C 549	14.2	71.0	2082	13	US-10-027-632-98674	Sequence 98674, A	C 622	14.2	71.0	3265	9	US-09-993-604-378	Sequence 378, App
C 550	14.2	71.0	2082	16	US-10-027-632-98674	Sequence 98674, A	C 623	14.2	71.0	3265	9	US-09-990-456-378	Sequence 378, App
C 551	14.2	71.0	2109	9	US-09-925-297-83	Sequence 83, App1	C 624	14.2	71.0	3265	9	US-09-989-721-378	Sequence 378, App
C 552	14.2	71.0	2110	16	US-10-104-047-743	Sequence 743, App	C 625	14.2	71.0	3265	9	US-09-992-558-378	Sequence 378, App
C 553	14.2	71.0	2194	17	US-10-437-963-62181	Sequence 62181, A	C 626	14.2	71.0	3265	9	US-09-989-253A-378	Sequence 378, App
C 554	14.2	71.0	2200	17	US-09-925-301-217	Sequence 217, App	C 627	14.2	71.0	3265	9	US-09-989-735-378	Sequence 378, App
C 555	14.2	71.0	2200	17	US-10-322-281-812	Sequence 812, App	C 628	14.2	71.0	3265	9	US-09-990-444-378	Sequence 378, App
C 556	14.2	71.0	2211	15	US-10-171-581-119	Sequence 119, App	C 629	14.2	71.0	3265	9	US-09-991-181-378	Sequence 378, App
C 557	14.2	71.0	2215	13	US-10-283-122A-33208	Sequence 33208, A	C 630	14.2	71.0	3265	9	US-09-989-730-378	Sequence 378, App
C 558	14.2	71.0	2265	17	US-10-437-963-49826	Sequence 49826, A	C 631	14.2	71.0	3265	9	US-09-990-426-378	Sequence 378, App
C 559	14.2	71.0	2274	13	US-10-282-122A-31568	Sequence 31568, A	C 632	14.2	71.0	3265	9	US-09-993-687-378	Sequence 378, App
C 560	14.2	71.0	2313	16	US-10-369-493-41968	Sequence 41968, A	C 633	14.2	71.0	3265	10	US-09-989-734-378	Sequence 378, App
C 561	14.2	71.0	2344	15	US-10-352-762-35	Sequence 35, App1	C 634	14.2	71.0	3265	10	US-09-997-653-378	Sequence 378, App
C 562	14.2	71.0	2347	15	US-10-352-762-1	Sequence 1, App1	C 635	14.2	71.0	3265	10	US-09-993-667-378	Sequence 378, App
C 563	14.2	71.0	2347	15	US-10-352-762-34	Sequence 34, App1	C 636	14.2	71.0	3265	10	US-09-997-428-378	Sequence 378, App
C 564	14.2	71.0	2347	15	US-10-352-762-36	Sequence 36, App1	C 637	14.2	71.0	3265	10	US-09-997-666-378	Sequence 378, App
C 565	14.2	71.0	2373	9	US-09-964-824A-253	Sequence 253, App	C 638	14.2	71.0	3265	10	US-09-990-437-378	Sequence 378, App
C 566	14.2	71.0	2373	9	US-09-880-107-2365	Sequence 2365, Ap	C 639	14.2	71.0	3265	10	US-09-990-443-378	Sequence 378, App
C 567	14.2	71.0	2373	15	US-10-101-510-745	Sequence 745, App	C 640	14.2	71.0	3265	10	US-09-990-711-378	Sequence 378, App
C 568	14.2	71.0	2373	13	US-10-641-643-1011	Sequence 1011, Ap	C 641	14.2	71.0	3265	10	US-09-989-126-378	Sequence 378, App
C 569	14.2	71.0	2429	14	US-10-116-802-98	Sequence 98, App1	C 642	14.2	71.0	3265	10	US-09-990-437-378	Sequence 378, App
C 570	14.2	71.0	2429	14	US-10-044-090-344	Sequence 344, App	C 643	14.2	71.0	3265	10	US-09-991-157-378	Sequence 378, App
C 571	14.2	71.0	2436	17	US-10-437-963-283	Sequence 283, App	C 644	14.2	71.0	3265	10	US-09-997-514-378	Sequence 378, App
C 572	14.2	71.0	2485	15	US-10-084-817-228	Sequence 228, App	C 645	14.2	71.0	3265	10	US-09-997-514-378	Sequence 378, App
C 573	14.2	71.0	2485	15	US-10-101-510-508	Sequence 508, App	C 646	14.2	71.0	3265	10	US-09-997-514-378	Sequence 378, App
C 574	14.2	71.0	2485	15	US-10-253-157-206	Sequence 206, App	C 647	14.2	71.0	3265	10	US-09-991-172-378	Sequence 378, App
C 575	14.2	71.0	2487	13	US-10-283-132A-32996	Sequence 32996, A	C 648	14.2	71.0	3265	10	US-09-990-726-378	Sequence 378, App
C 576	14.2	71.0	2502	17	US-10-437-963-95565	Sequence 95565, A	C 649	14.2	71.0	3265	10	US-09-997-559-378	Sequence 378, App
C 577	14.2	71.0	2506	12	US-10-041-018-377	Sequence 377, App	C 650	14.2	71.0	3265	10	US-09-997-601-378	Sequence 378, App
C 578	14.2	71.0	2562	16	US-10-369-493-48567	Sequence 48567, A	C 651	14.2	71.0	3265	10	US-09-990-443-378	Sequence 378, App
C 579	14.2	71.0	2574	17	US-10-437-963-34290	Sequence 34290, A	C 652	14.2	71.0	3265	10	US-09-991-854-378	Sequence 378, App
C 580	14.2	71.0	2607	16	US-10-369-493-31325	Sequence 31325, A	C 653	14.2	71.0	3265	10	US-09-997-628-378	Sequence 378, App
C 581	14.2	71.0	2684	17	US-10-283-975A-48	Sequence 48, App1	C 654	14.2	71.0	3265	10	US-09-987-683-378	Sequence 378, App
C 582	14.2	71.0	2754	15	US-10-345-680-33	Sequence 33, App1	C 655	14.2	71.0	3265	10	US-09-989-729A-378	Sequence 378, App
C 583	14.2	71.0	2771	15	US-10-223-567A-468	Sequence 468, App	C 656	14.2	71.0	3265	10	US-09-997-440-378	Sequence 378, App
C 584	14.2	71.0	2825	13	US-10-425-114-42222	Sequence 22222, A	C 657	14.2	71.0	3265	10	US-09-997-440-378	Sequence 378, App
C 585	14.2	71.0	2834	9	US-09-954-456-286	Sequence 286, App	C 658	14.2	71.0	3265	10	US-09-990-440-378	Sequence 378, App
C 586	14.2	71.0	2834	9	US-09-954-456-951	Sequence 951, App	C 659	14.2	71.0	3265	10	US-09-993-469-378	Sequence 378, App
C 587	14.2	71.0	2834	9	US-09-954-456-1603	Sequence 1603, App	C 660	14.2	71.0	3265	10	US-09-997-542-378	Sequence 378, App
C 588	14.2	71.0	2834	16	US-09-873-367C-311	Sequence 311, App	C 661	14.2	71.0	3265	10	US-09-993-748-378	Sequence 378, App
C 589	14.2	71.0	2834	16	US-10-295-027-874	Sequence 874, App	C 662	14.2	71.0	3265	10	US-09-990-439-378	Sequence 378, App
C 590	14.2	71.0	2841	15	US-10-128-714-57	Sequence 57, App1	C 663	14.2	71.0	3265	10	US-09-990-427-378	Sequence 378, App
C 591	14.2	71.0	2841	15	US-10-128-714-5057	Sequence 5057, App	C 664	14.2	71.0	3265	10	US-09-989-328-378	Sequence 378, App
C 592	14.2	71.0	2852	16	US-10-094-749-1113	Sequence 1113, App	C 665	14.2	71.0	3265	10	US-09-993-583-378	Sequence 378, App
C 593	14.2	71.0	2867	13	US-10-116-802-155	Sequence 155, App	C 666	14.2	71.0	3265	10	US-09-941-892-378	Sequence 378, App
C 594	14.2	71.0	2981	15	US-10-096-534-45	Sequence 45, App1	C 667	14.2	71.0	3265	10	US-09-992-521-378	Sequence 378, App
C 595	14.2	71.0	3003	16	US-10-429-873A-5	Sequence 5, App1	C 668	14.2	71.0	3265	10	US-09-997-333-378	Sequence 378, App
C 596	14.2	71.0	3003	16	US-10-429-873A-6	Sequence 6, App1	C 669	14.2	71.0	3265	10	US-09-997-384-378	Sequence 378, App
C 597	14.2	71.0	3043	14	US-10-025-167-16	Sequence 16, App1	C 670	14.2	71.0	3265	10	US-09-998-041-378	Sequence 378, App
C 598	14.2	71.0	3044	15	US-10-233-665-38	Sequence 38, App1	C 671	14.2	71.0	3265	10	US-09-997-585-378	Sequence 378, App

C 672	14.2	71.0	3265	10	US-09-997-614-378	Sequence 378, App
C 673	14.2	71.0	3265	10	US-09-989-862-378	Sequence 378, App
C 674	14.2	71.0	3265	10	US-09-997-529-378	Sequence 378, App
C 675	14.2	71.0	3265	10	US-09-989-725-378	Sequence 378, App
C 676	14.2	71.0	3265	11	US-09-989-733-378	Sequence 378, App
C 677	14.2	71.0	3265	11	US-09-992-643-378	Sequence 378, App
C 678	14.2	71.0	3265	13	US-10-206-915-257	Sequence 257, App
C 679	14.2	71.0	3265	13	US-10-199-670-257	Sequence 257, App
C 680	14.2	71.0	3265	13	US-10-201-858-257	Sequence 257, App
C 681	14.2	71.0	3265	13	US-10-205-890-257	Sequence 257, App
C 682	14.2	71.0	3265	13	US-10-208-024-257	Sequence 257, App
C 683	14.2	71.0	3265	13	US-10-201-853-257	Sequence 257, App
C 684	14.2	71.0	3265	13	US-10-063-745-69	Sequence 69, App1
C 685	14.2	71.0	3265	13	US-09-989-724-378	Sequence 378, App
C 686	14.2	71.0	3265	13	US-09-989-728-378	Sequence 378, App
C 687	14.2	71.0	3265	13	US-09-990-441-378	Sequence 378, App
C 688	14.2	71.0	3265	13	US-10-063-512-69	Sequence 69, App1
C 689	14.2	71.0	3265	13	US-10-063-513-69	Sequence 69, App1
C 690	14.2	71.0	3265	13	US-10-063-515-69	Sequence 69, App1
C 691	14.2	71.0	3265	13	US-10-063-549-69	Sequence 69, App1
C 692	14.2	71.0	3265	13	US-10-063-569-69	Sequence 69, App1
C 693	14.2	71.0	3265	13	US-10-063-551-69	Sequence 69, App1
C 694	14.2	71.0	3265	13	US-10-174-581-257	Sequence 257, App
C 695	14.2	71.0	3265	13	US-10-176-483-257	Sequence 257, App
C 696	14.2	71.0	3265	13	US-10-176-749-257	Sequence 257, App
C 697	14.2	71.0	3265	13	US-10-176-914-257	Sequence 257, App
C 698	14.2	71.0	3265	13	US-10-176-915-257	Sequence 257, App
C 699	14.2	71.0	3265	13	US-09-997-857-378	Sequence 378, App
C 700	14.2	71.0	3265	13	US-10-063-555-69	Sequence 69, App1
C 701	14.2	71.0	3265	13	US-10-063-563-69	Sequence 69, App1
C 702	14.2	71.0	3265	13	US-10-063-594-69	Sequence 69, App1
C 703	14.2	71.0	3265	13	US-10-063-553-69	Sequence 69, App1
C 704	14.2	71.0	3265	13	US-10-063-554-69	Sequence 69, App1
C 705	14.2	71.0	3265	13	US-10-176-484-257	Sequence 257, App
C 706	14.2	71.0	3265	13	US-10-180-550-257	Sequence 257, App
C 707	14.2	71.0	3265	13	US-10-183-014-257	Sequence 257, App
C 708	14.2	71.0	3265	13	US-10-187-738-257	Sequence 257, App
C 709	14.2	71.0	3265	13	US-10-187-740-257	Sequence 257, App
C 710	14.2	71.0	3265	13	US-10-187-883-257	Sequence 257, App
C 711	14.2	71.0	3265	13	US-10-194-363-257	Sequence 257, App
C 712	14.2	71.0	3265	13	US-10-194-460-257	Sequence 257, App
C 713	14.2	71.0	3265	13	US-10-194-463-257	Sequence 257, App
C 714	14.2	71.0	3265	13	US-10-194-484-257	Sequence 257, App
C 715	14.2	71.0	3265	13	US-10-195-884-257	Sequence 257, App
C 716	14.2	71.0	3265	13	US-10-195-896-257	Sequence 257, App
C 717	14.2	71.0	3265	13	US-10-196-744-257	Sequence 257, App
C 718	14.2	71.0	3265	13	US-10-196-755-257	Sequence 257, App
C 719	14.2	71.0	3265	13	US-10-196-757-257	Sequence 257, App
C 720	14.2	71.0	3265	13	US-10-197-704-257	Sequence 257, App
C 721	14.2	71.0	3265	13	US-10-197-710-257	Sequence 257, App
C 722	14.2	71.0	3265	13	US-10-198-758-257	Sequence 257, App
C 723	14.2	71.0	3265	13	US-10-198-766-257	Sequence 257, App
C 724	14.2	71.0	3265	13	US-10-199-304-257	Sequence 257, App
C 725	14.2	71.0	3265	13	US-10-199-309-257	Sequence 257, App
C 726	14.2	71.0	3265	13	US-10-199-313-257	Sequence 257, App
C 727	14.2	71.0	3265	13	US-10-199-456-257	Sequence 257, App
C 728	14.2	71.0	3265	13	US-10-201-329-257	Sequence 257, App
C 729	14.2	71.0	3265	13	US-10-202-412-257	Sequence 257, App
C 730	14.2	71.0	3265	13	US-10-206-919-257	Sequence 257, App
C 731	14.2	71.0	3265	13	US-10-206-922-257	Sequence 257, App
C 732	14.2	71.0	3265	13	US-10-206-924-257	Sequence 257, App
C 733	14.2	71.0	3265	13	US-10-206-928-257	Sequence 257, App
C 734	14.2	71.0	3265	13	US-10-207-914-257	Sequence 257, App
C 735	14.2	71.0	3265	13	US-10-207-921-257	Sequence 257, App
C 736	14.2	71.0	3265	13	US-10-207-922-257	Sequence 257, App
C 737	14.2	71.0	3265	13	US-10-208-027-257	Sequence 257, App
C 738	14.2	71.0	3265	13	US-09-997-641-378	Sequence 378, App
C 739	14.2	71.0	3265	13	US-09-991-150-378	Sequence 378, App
C 740	14.2	71.0	3265	13	US-10-174-570-257	Sequence 257, App
C 741	14.2	71.0	3265	13	US-10-183-005-257	Sequence 257, App
C 742	14.2	71.0	3265	14	US-10-006-867-69	Sequence 69, App1
C 743	14.2	71.0	3265	14	US-10-052-586-257	Sequence 257, App
C 744	14.2	71.0	3265	14	US-10-063-547-69	Sequence 69, App1
C 745	14.2	71.0	3265	15	US-10-174-590-257	Sequence 257, App
C 746	14.2	71.0	3265	15	US-10-176-758-257	Sequence 257, App
C 747	14.2	71.0	3265	15	US-10-175-737-257	Sequence 257, App
C 748	14.2	71.0	3265	15	US-10-063-616-69	Sequence 69, App1
C 749	14.2	71.0	3265	15	US-10-173-706-257	Sequence 257, App
C 750	14.2	71.0	3265	15	US-10-175-738-257	Sequence 257, App
C 751	14.2	71.0	3265	15	US-10-175-752-257	Sequence 257, App
C 752	14.2	71.0	3265	15	US-10-176-482-257	Sequence 257, App
C 753	14.2	71.0	3265	15	US-10-176-757-257	Sequence 257, App
C 754	14.2	71.0	3265	15	US-10-176-913-257	Sequence 257, App
C 755	14.2	71.0	3265	15	US-10-180-552-257	Sequence 257, App
C 756	14.2	71.0	3265	15	US-10-180-557-257	Sequence 257, App
C 757	14.2	71.0	3265	15	US-10-063-802-69	Sequence 69, App1
C 758	14.2	71.0	3265	15	US-10-173-700-257	Sequence 257, App
C 759	14.2	71.0	3265	15	US-10-174-572-257	Sequence 257, App
C 760	14.2	71.0	3265	15	US-10-174-579-257	Sequence 257, App
C 761	14.2	71.0	3265	15	US-10-174-582-257	Sequence 257, App
C 762	14.2	71.0	3265	15	US-10-174-588-257	Sequence 257, App
C 763	14.2	71.0	3265	15	US-10-175-739-257	Sequence 257, App
C 764	14.2	71.0	3265	15	US-10-175-740-257	Sequence 257, App
C 765	14.2	71.0	3265	15	US-10-175-743-257	Sequence 257, App
C 766	14.2	71.0	3265	15	US-10-176-488-257	Sequence 257, App
C 767	14.2	71.0	3265	15	US-10-176-492-257	Sequence 257, App
C 768	14.2	71.0	3265	15	US-10-176-747-257	Sequence 257, App
C 769	14.2	71.0	3265	15	US-10-176-750-257	Sequence 257, App
C 770	14.2	71.0	3265	15	US-10-176-985-257	Sequence 257, App
C 771	14.2	71.0	3265	15	US-10-176-987-257	Sequence 257, App
C 772	14.2	71.0	3265	15	US-10-176-992-257	Sequence 257, App
C 773	14.2	71.0	3265	15	US-10-176-993-257	Sequence 257, App
C 774	14.2	71.0	3265	15	US-10-184-658-257	Sequence 257, App
C 775	14.2	71.0	3265	15	US-10-176-991-257	Sequence 257, App
C 776	14.2	71.0	3265	15	US-10-173-695-257	Sequence 257, App
C 777	14.2	71.0	3265	15	US-10-173-697-257	Sequence 257, App
C 778	14.2	71.0	3265	15	US-10-173-705-257	Sequence 257, App
C 779	14.2	71.0	3265	15	US-10-174-576-257	Sequence 257, App
C 780	14.2	71.0	3265	15	US-10-174-585-257	Sequence 257, App
C 781	14.2	71.0	3265	15	US-10-174-586-257	Sequence 257, App
C 782	14.2	71.0	3265	15	US-10-175-747-257	Sequence 257, App
C 783	14.2	71.0	3265	15	US-10-176-481-257	Sequence 257, App
C 784	14.2	71.0	3265	15	US-10-176-485-257	Sequence 257, App
C 785	14.2	71.0	3265	15	US-10-176-487-257	Sequence 257, App
C 786	14.2	71.0	3265	15	US-10-176-493-257	Sequence 257, App
C 787	14.2	71.0	3265	15	US-10-176-756-257	Sequence 257, App
C 788	14.2	71.0	3265	15	US-10-176-911-257	Sequence 257, App
C 789	14.2	71.0	3265	15	US-10-176-919-257	Sequence 257, App
C 790	14.2	71.0	3265	15	US-10-176-925-257	Sequence 257, App
C 791	14.2	71.0	3265	15	US-10-179-510-257	Sequence 257, App
C 792	14.2	71.0	3265	15	US-10-179-510-257	Sequence 257, App
C 793	14.2	71.0	3265	15	US-10-180-543-257	Sequence 257, App
C 794	14.2	71.0	3265	15	US-10-180-544-257	Sequence 257, App
C 795	14.2	71.0	3265	15	US-10-180-546-257	Sequence 257, App
C 796	14.2	71.0	3265	15	US-10-180-547-257	Sequence 257, App
C 797	14.2	71.0	3265	15	US-10-180-549-257	Sequence 257, App
C 798	14.2	71.0	3265	15	US-10-180-555-257	Sequence 257, App
C 799	14.2	71.0	3265	15	US-10-180-559-257	Sequence 257, App
C 800	14.2	71.0	3265	15	US-10-181-000-257	Sequence 257, App
ALIGNMENTS						
RESULT 1						
US-09-825-489-3						
; Sequence 3, Application US/09825489						
; Publication No. US20030232767A1						
; GENERAL INFORMATION:						
; APPLICANT: AGRAWAL, SUDHIR						
; APPLICANT: KANDIMALLA, EKAMBAR R.						
; APPLICANT: BREGMAN, DAVID B.						
; APPLICANT: MANT, SRIDHAR						
; APPLICANT: LU, YI						
; TITLE OF INVENTION: SENSITIZATION OF CELLS TO CYTOTOXIC AGENTS USING						
; TITLE OF INVENTION: OLIGONUCLEOTIDES DIRECTED TO NUCLEOTIDE EXCISION REPAIR						

```

; TITLE OF INVENTION: OR TRANSCRIPTION COUPLED REPAIR GENES
; FILE REFERENCE: HYZ-075US2 (475.08.514)
; CURRENT APPLICATION NUMBER: US/09/825,489
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-825-489-3
```

```

Query Match          100.0%; Score 20; DB 11; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCTCATGTTGATG 20
        |||||
Db       1 GGTCCATCTCATGTTGATG 20
```

```

RESULT 2
US-09-867-701-7544
; Sequence 7544, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7544
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(486)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-7544
```

```

Query Match          100.0%; Score 20; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCTCATGTTGATG 20
        |||||
Db       149 GGTCCATCTCATGTTGATG 168
```

```

RESULT 3
US-09-867-701-2117/c
; Sequence 2117, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 2117
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(506)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2117
```

```

Query Match          100.0%; Score 20; DB 9; Length 506;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCTCATGTTGATG 20
        |||||
Db       246 GGTCCATCTCATGTTGATG 227
```

```

RESULT 4
US-09-867-701-2107
; Sequence 2107, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2107
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(513)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2107
```

```

Query Match          100.0%; Score 20; DB 9; Length 513;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GGTCCATCTCATGTTGATG 20
        |||||
Db       357 GGTCCATCTCATGTTGATG 376
```

```

RESULT 5
US-09-825-489-13/c
; Sequence 13, Application US/09825489
; Publication No. US20030232767A1
; GENERAL INFORMATION:
; APPLICANT: AGRAMAL, SUDHIR
; APPLICANT: KANDIMALLA, EKAMBAR R.
; APPLICANT: BREGMAN, DAVID B.
; APPLICANT: MANI, SRIDHAR
; APPLICANT: LU, YI
; TITLE OF INVENTION: SENSITIZATION OF CELLS TO CYTOTOXIC AGENTS USING
; FILE REFERENCE: HYZ-075US2 (475.08.514)
; CURRENT APPLICATION NUMBER: US/09/825,489
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1377
```

/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-09-825-489-13

Query Match 100.0%; Score 20; DB 11; Length 1377;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
DB 769 GGTCCATCTCATGTTGATG 750

RESULT 6  
US-10-103-313-116/c  
; Sequence 116, Application US/10103313  
; Publication No. US20030082758A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PJ207C1  
; CURRENT APPLICATION NUMBER: US/10/103,313  
; CURRENT FILING DATE: 2002-03-12  
; NUMBER OF SEQ ID NOS: 653  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 116  
; LENGTH: 1407  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-103-313-116

Query Match 100.0%; Score 20; DB 15; Length 1407;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
DB 782 GGTCCATCTCATGTTGATG 763

RESULT 7  
US-10-101-510-587  
; Sequence 587, Application US/10101510  
; Publication No. US20030148295A1  
; GENERAL INFORMATION:  
; APPLICANT: WAN, JACKSON  
; APPLICANT: WANG, YIXIN  
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
; FILE REFERENCE: 15117.0012  
; CURRENT APPLICATION NUMBER: US/10/101,510  
; CURRENT FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/276,947  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 805  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 587  
; LENGTH: 4670  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (1781)..(1894)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
; NAME/KEY: modified\_base  
; LOCATION: (2658)..(3015)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-101-510-587

Query Match 100.0%; Score 20; DB 15; Length 4670;  
Best Local Similarity 100.0%; Pred. No. 6.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
DB 695 GGTCCATCTCATGTTGATG 714

RESULT 8  
US-10-027-632-257254/c  
; Sequence 257254, Application US/10027632  
; Publication No. US20020196371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 257254  
; LENGTH: 1526  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-257254

Query Match 85.0%; Score 17; DB 13; Length 1526;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CCATCTCATGTTGATG 20  
|||||  
DB 926 CCATCTCATGTTGATG 910

RESULT 9  
US-10-027-632-257254/c  
; Sequence 257254, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09



NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 257254  
LENGTH: 1526  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-257254

Query Match  
Best Local Similarity 85.0%; Score 17; DB 16; Length 1526;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTTGATG 20  
Db 926 CCATCTCATGTTGATG 910

RESULT 10  
US-10-158-844-40/C  
Sequence 40, Application US/10158844  
Publication No. US20040029118A1  
GENERAL INFORMATION:

APPLICANT: Kunsch et al.  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-R  
COMPUTER: Dell Latitude Pentium 3  
OPERATING SYSTEM: Windows 98  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/158,844  
FILING DATE: 03-Jun-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/961,527  
FILING DATE: 1997-10-30  
APPLICATION NUMBER: US 60/029,960  
FILING DATE: 1996-10-31

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB340P1D1

INFORMATION FOR SEQ ID NO: 40:

SEQUENCE CHARACTERISTICS:

LENGTH: 14273 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 40:

US-10-158-844-40

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20  
Db 13385 GGTCCATCTCATGTTGATG 13366

RESULT 11  
US-10-027-632-253417/C  
Sequence 253417, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 253417  
LENGTH: 1049  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-253417

Query Match  
Best Local Similarity 82.0%; Score 16.4; DB 13;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTTGATG 20  
Db 71 TCCATCTCATGTTGATG 54

RESULT 12

US-10-027-632-253417/C  
Sequence 253417, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 253417

LENGTH: 1049

TYPE: DNA

ORGANISM: Human

US-10-027-632-253417

Query Match  
Best Local Similarity 82.0%; Score 16.4; DB 16;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGATG 20  
|||||  
DB 71 TCCATTCATCATGTTGATG 54

RESULT 13  
US-10-085-783A-40734  
/ Sequence 40734, Application US/10085783A  
/ Publication No. US20040037841A1  
/ GENERAL INFORMATION:  
/ APPLICANT: ChondroGene Inc.  
/ APPLICANT: Liew, C.C.  
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
/ FILE REFERENCE: 4231/2002  
/ CURRENT APPLICATION NUMBER: US/10/085,783A  
/ PRIOR FILING DATE: 2002-02-28  
/ PRIOR APPLICATION NUMBER: US 60/305,340  
/ PRIOR FILING DATE: 2001-07-13  
/ PRIOR APPLICATION NUMBER: US 60/275,017  
/ PRIOR FILING DATE: 2001-03-12  
/ PRIOR APPLICATION NUMBER: US 60/271,955  
/ PRIOR FILING DATE: 2001-02-28  
/ NUMBER OF SEQ ID NOS: 58994  
/ SOFTWARE: PatentIn version 3.2  
/ SEQ ID NO 40734  
/ LENGTH: 238  
/ TYPE: DNA  
/ ORGANISM: Human  
US-10-085-783A-40734

Query Match 79.0%; Score 15.8; DB 13; Length 238;  
Best Local Similarity 89.5%; Pred. No. 5.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20  
|||||  
DB 77 GTCCATATCAATCAAGTTGATG 95

RESULT 14  
US-10-242-535A-40734  
/ Sequence 40734, Application US/10242535A  
/ Publication No. US20040013663A1  
/ GENERAL INFORMATION:  
/ APPLICANT: ChondroGene Inc.  
/ APPLICANT: Liew, C.C.  
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
/ FILE REFERENCE: 4231/2005  
/ CURRENT APPLICATION NUMBER: US/10/242,535A  
/ PRIOR FILING DATE: 2002-09-12  
/ PRIOR APPLICATION NUMBER: US 10/085,783  
/ PRIOR FILING DATE: 2002-02-28  
/ PRIOR APPLICATION NUMBER: US 60/305,340  
/ PRIOR FILING DATE: 2001-07-13  
/ PRIOR APPLICATION NUMBER: US 60/275,017  
/ PRIOR FILING DATE: 2001-03-12  
/ PRIOR APPLICATION NUMBER: US 60/271,955  
/ PRIOR FILING DATE: 2001-02-28  
/ NUMBER OF SEQ ID NOS: 58994  
/ SOFTWARE: PatentIn version 3.2  
/ SEQ ID NO 40734  
/ LENGTH: 238  
/ TYPE: DNA  
/ ORGANISM: Human  
US-10-242-535A-40734

Query Match 79.0%; Score 15.8; DB 16; Length 238;  
Best Local Similarity 89.5%; Pred. No. 5.2e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20  
|||||

DB 77 GTCCATATCAATCAAGTTGATG 95

RESULT 15  
US-09-732-627A-2949/C  
/ Sequence 2949, Application US/09732627A  
/ Publication No. US20040123338A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Fincher, Karen L.  
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
/ FILE REFERENCE: 38-21(51770)B  
/ CURRENT APPLICATION NUMBER: US/09/732,627A  
/ CURRENT FILING DATE: 2000-12-08  
/ NUMBER OF SEQ ID NOS: 4930  
/ SEQ ID NO 2949  
/ LENGTH: 296  
/ TYPE: DNA  
/ ORGANISM: Gossypium hirsutum  
/ FEATURE:  
/ OTHER INFORMATION: Clone ID: LIB3493-010-Pl-M1-E9  
US-09-732-627A-2949

Query Match 79.0%; Score 15.8; DB 12; Length 296;  
Best Local Similarity 89.5%; Pred. No. 5.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20  
|||||  
DB 29 GTCCATACATCATGTTGATG 11

RESULT 16  
US-10-060-036-3118/C  
/ Sequence 3118, Application US/10060036  
/ Publication No. US20030073144A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Benson, Darin R.  
/ APPLICANT: Kalos, Michael D.  
/ APPLICANT: Lodes, Michael J.  
/ APPLICANT: Persing, David H.  
/ APPLICANT: Hepler, William T.  
/ APPLICANT: Jjiang, Yugu  
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
/ FILE REFERENCE: 210121.566  
/ CURRENT APPLICATION NUMBER: US/10/060,036  
/ CURRENT FILING DATE: 2002-01-30  
/ NUMBER OF SEQ ID NOS: 4560  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 3118  
/ LENGTH: 431  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: misc feature  
/ LOCATION: 111..284  
/ OTHER INFORMATION: n = A,T,C or G  
US-10-060-036-3118

Query Match 79.0%; Score 15.8; DB 15; Length 431;  
Best Local Similarity 89.5%; Pred. No. 5.8e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATACATCATGTTGATG 20  
|||||  
DB 194 GTCCATATCAATCAAGTTGATG 176

RESULT 17  
US-09-918-995-12216/C  
/ Sequence 12216, Application US/09918995  
/ Publication No. US20030073623A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
/ TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
/ FILE REFERENCE: 20411-756
/ CURRENT APPLICATION NUMBER: US/09/918,995
/ CURRENT FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: US/09/235,076
/ PRIOR FILING DATE: 1999-01-20
/ NUMBER OF SEQ ID NOS: 38054
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 12216
/ LENGTH: 462
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(462)
/ OTHER INFORMATION: n = A,T,C or G
US-09-918-995-12216

Query Match          79.0%; Score 15.8; DB 10; Length 462;
Best Local Similarity 89.5%; Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GTCCATCTCATGTTGATG 20
Db      446  GTCCCTCTCATGTTGATG 428

RESULT 18
US-10-085-783A-45577/C
/ Sequence 45577, Application US/10085783A
/ Publication No. US20040037841A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2002
/ CURRENT APPLICATION NUMBER: US/10/085,783A
/ CURRENT FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 45577
/ LENGTH: 540
/ TYPE: DNA
/ ORGANISM: Human
US-10-085-783A-45577

Query Match          79.0%; Score 15.8; DB 13; Length 540;
Best Local Similarity 89.5%; Pred. No. 6.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GGTCATCTCATGTTGAT 19
Db      233  GGTCACATCATGTTGAT 215

RESULT 19
US-10-242-535A-45577/C
/ Sequence 45577, Application US/10242535A
/ Publication No. US20040013663A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2005
```

```
/ CURRENT APPLICATION NUMBER: US/10/242,535A
/ CURRENT FILING DATE: 2002-09-12
/ PRIOR APPLICATION NUMBER: US 10/085,783
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 45577
/ LENGTH: 540
/ TYPE: DNA
/ ORGANISM: Human
US-10-242-535A-45577

Query Match          79.0%; Score 15.8; DB 16; Length 540;
Best Local Similarity 89.5%; Pred. No. 6.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  GTCCATCTCATGTTGAT 19
Db      233  GTCCACATCATGTTGAT 215

RESULT 20
US-09-871-161-62
/ Sequence 62, Application US/09871161
/ Publication No. US20030097666A1
/ GENERAL INFORMATION:
/ APPLICANT: ENDBERG, WILSON O., ET AL.
/ TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
/ FILE REFERENCE: CCDNA-260XX
/ CURRENT APPLICATION NUMBER: US/09/871,161
/ CURRENT FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 09/328,111
/ PRIOR FILING DATE: 1999-06-08
/ PRIOR APPLICATION NUMBER: 60/117,393
/ PRIOR FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: 60/098,639
/ PRIOR FILING DATE: 1998-08-31
/ NUMBER OF SEQ ID NOS: 544
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 62
/ LENGTH: 614
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(614)
/ OTHER INFORMATION: n = A,T,C or G
US-09-871-161-62

Query Match          79.0%; Score 15.8; DB 10; Length 614;
Best Local Similarity 89.5%; Pred. No. 6.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2  GTCCATCTCATGTTGAT 20
Db      466  GTCCATATCAAGTTGATG 484

RESULT 21
US-10-641-643-1225/C
/ Sequence 1225, Application US/10641643
/ Publication No. US20040077003A1
/ GENERAL INFORMATION:
/ APPLICANT: Cocks, Benjamin G.
/ APPLICANT: Susan G. Stuart
/ APPLICANT: Jeffrey J. Seilhamer
```

```
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
/ GENE EXPRESSION
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/641,643
/ FILING DATE: 14-Aug-2003
/ CLASSIFICATION: <Unknown>
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: <Unknown>
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1225:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 686 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g246741
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1225 :
/
US-10-641-643-1225
/
Query Match          79.0%; Score 15.8; DB 17; Length 686;
Best Local Similarity 89.5%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
QY      1  GTCCATCTCATGTTGAT 19
DB      628  GTCCATCTCATGTTGAT 610
/
RESULT 22
US-10-424-599-20394
/ Sequence 20394, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 20394
/ LENGTH: 765
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_118420C.1
/
US-10-424-599-20394
```

```
Query Match          79.0%; Score 15.8; DB 13; Length 765;
Best Local Similarity 89.5%; Pred. No. 6.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
QY      2  GTCCATCTCATGTTGATG 20
DB      484  GTCCATCTCATGTTGATG 502
/
RESULT 23
US-10-437-963-64475
/ Sequence 64475, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 64475
/ LENGTH: 789
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)-(789)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_65615C.1
/
US-10-437-963-64475
/
Query Match          79.0%; Score 15.8; DB 17; Length 789;
Best Local Similarity 89.5%; Pred. No. 6.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
QY      2  GTCCATCTCATGTTGATG 20
DB      442  GTCCATCTCATGTTGATG 460
/
RESULT 24
US-10-437-963-7531
/ Sequence 7531, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 7531
/ LENGTH: 1329
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
```

OTHER INFORMATION: Clone ID: PAT\_MRT4530\_14117C.1  
US-10-437-963-7531

Query Match 79.0%; Score 15.8; DB 17; Length 1329;  
Best Local Similarity 89.5%; Pred. No. 7.1e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20  
DB 355 GTCCATCTCATGTTGATG 373

RESULT 25  
US-10-041-018-111

Sequence 111, Application US/10041018  
Publication No. US20040072323A1  
GENERAL INFORMATION:  
APPLICANT: Matsuda, Seichi P.T.  
APPLICANT: Hart, Elizabeth A.  
TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
FILE REFERENCE: P02080US1/10025547  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259880  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 413  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 111  
LENGTH: 3021  
TYPE: DNA  
ORGANISM: Dendroctonus jeffreyi  
US-10-041-018-111

Query Match 79.0%; Score 15.8; DB 12; Length 3021;  
Best Local Similarity 89.5%; Pred. No. 8.3e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20  
DB 2142 GTCCATCTCATGTTGATG 2160

RESULT 26  
US-10-041-018-112

Sequence 112, Application US/10041018  
Publication No. US20040072323A1  
GENERAL INFORMATION:  
APPLICANT: Matsuda, Seichi P.T.  
APPLICANT: Hart, Elizabeth A.  
TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
FILE REFERENCE: P02080US1/10025547  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259880  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 413  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 112  
LENGTH: 5874  
TYPE: DNA  
ORGANISM: Dendroctonus jeffreyi  
US-10-041-018-112

Query Match 79.0%; Score 15.8; DB 12; Length 5874;  
Best Local Similarity 89.5%; Pred. No. 9.4e+02;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20  
DB 4943 GTCCATCTCATGTTGATG 4961

RESULT 27

US-10-087-192-460  
Sequence 460, Application US/10087192  
Publication No. US20020182586A1  
GENERAL INFORMATION:  
APPLICANT: Morris, David W.  
APPLICANT: Engelhard, Eric K.  
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
FILE REFERENCE: 529452000122  
CURRENT APPLICATION NUMBER: US/10/087,192  
CURRENT FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: US 09/747,377  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 09/798,586  
PRIOR FILING DATE: 2001-03-02  
NUMBER OF SEQ ID NOS: 2059  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 460  
LENGTH: 101209  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(101209)  
OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-460

Query Match 79.0%; Score 15.8; DB 13; Length 101209;  
Best Local Similarity 89.5%; Pred. No. 1.5e+03;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20  
DB 39900 GTCCATCTCATGTTGATG 39918

RESULT 28  
US-10-027-632-79848/C

Sequence 79848, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 79848  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-79848

Query Match 77.0%; Score 15.4; DB 13; Length 519;  
Best Local Similarity 94.1%; Pred. No. 9.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      4 CCATCTCATGTTGATG 20
      |||||
      253 CCATACCCATGTTGATG 237

RESULT 29
US-10-027-632-79849/c
; Sequence 79849, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79849
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-79849

Query Match      77.0%; Score 15.4; DB 13; Length 519;
Best Local Similarity 94.1%; Pred. No. 9.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 CCATCTCATGTTGATG 20
      |||||
      253 CCATACCCATGTTGATG 237

Db      CCATACCCATGTTGATG 237

RESULT 30
US-10-027-632-79848/c
; Sequence 79848, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79848
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-79848

Query Match      77.0%; Score 15.4; DB 16; Length 519;
Best Local Similarity 94.1%; Pred. No. 9.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 CCATCTCATGTTGATG 20
      |||||
      253 CCATACCCATGTTGATG 237

Db      CCATACCCATGTTGATG 237

RESULT 31
US-10-027-632-79849/c
; Sequence 79849, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79849
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-79849

Query Match      77.0%; Score 15.4; DB 16; Length 519;
Best Local Similarity 94.1%; Pred. No. 9.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 CCATCTCATGTTGATG 20
      |||||
      253 CCATACCCATGTTGATG 237

Db      CCATACCCATGTTGATG 237

RESULT 32
US-09-764-847-497
; Sequence 497, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 497
; LENGTH: 551
```

```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (398)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (477)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (478)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (479)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: SITE
/ LOCATION: (532)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-764-847-497
```

```
Query Match          77.0%; Score 15.4; DB 9; Length 551;
Best Local Similarity 94.1%; Pred. No. 9.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      3  TCCATCTCATGTTGAT 19
         |||||
Db      261 TCCATCTCATGTTGAT 277
```

```
RESULT 33
US-10-092-154-497
/ Sequence 497, Application US/10092154
/ Publication No. US20030054375A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PC009C1
/ CURRENT APPLICATION NUMBER: US/10/092,154
/ NUMBER OF SEQ ID NOS: 2003
/ PRIORITY FILING DATE: 2002-03-07
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 497
/ LENGTH: 551
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (398)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (477)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (478)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (479)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (532)
/ OTHER INFORMATION: n equals a,t,g, or c
US-10-092-154-497
```

```
Query Match          77.0%; Score 15.4; DB 15; Length 551;
Best Local Similarity 94.1%; Pred. No. 9.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      3  TCCATCTCATGTTGAT 19
         |||||
Db      261 TCCATCTCATGTTGAT 277
```

```
RESULT 34
US-10-027-632-66707/c
/ Sequence 66707, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: Polymorphisms in the Human Genome
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIORITY FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 66707
/ LENGTH: 573
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-66707
```

```
Query Match          77.0%; Score 15.4; DB 13; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      3  TCCATCTCATGTTGAT 19
         |||||
Db      471 TCCATCTCATGTTGAT 455
```

```
RESULT 35
US-10-027-632-66708/c
/ Sequence 66708, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: Polymorphisms in the Human Genome
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIORITY FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 66708
/ LENGTH: 573
/ TYPE: DNA
/ ORGANISM: Human
```

US-10-027-632-66708

Query Match 77.0%; Score 15.4; DB 13; Length 573;  
Best Local Similarity 94.1%; Pred. No. 9.8e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19  
|||||  
Db 471 TCCATCTCATGTGTAT 455

RESULT 36

US-10-027-632-311197/c  
; Sequence 311196, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 311196  
; LENGTH: 573  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-311196

Query Match 77.0%; Score 15.4; DB 13; Length 573;  
Best Local Similarity 94.1%; Pred. No. 9.8e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19  
|||||  
Db 471 TCCATCTCATGTGTAT 455

RESULT 37

US-10-027-632-311197/c  
; Sequence 311197, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 311197  
; LENGTH: 573  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-311197

Query Match 77.0%; Score 15.4; DB 13; Length 573;  
Best Local Similarity 94.1%; Pred. No. 9.8e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19  
|||||  
Db 471 TCCATCTCATGTGTAT 455

RESULT 38

US-10-027-632-66707/c  
; Sequence 66707, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 66707  
; LENGTH: 573  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-66707

Query Match 77.0%; Score 15.4; DB 16; Length 573;  
Best Local Similarity 94.1%; Pred. No. 9.8e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 3 TCCATCTCATGTGTGAT 19  
|||||  
Db 471 TCCATCTCATGTGTAT 455

RESULT 39

US-10-027-632-66708/c  
; Sequence 66708, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129



```

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66708
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-66708

Query Match          77.0%; Score 15.4; DB 16; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTGAT 19
    |||||
Db 471 TCCATCTCATGTGAT 455

RESULT 40
US-10-027-632-311196/c
; Sequence 311196, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 311196
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-311196

Query Match          77.0%; Score 15.4; DB 16; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTGAT 19
    |||||
Db 471 TCCATCTCATGTGAT 455
```

```

RESULT 41
US-10-027-632-311197/c
; Sequence 311197, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 311197
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-311197

Query Match          77.0%; Score 15.4; DB 16; Length 573;
Best Local Similarity 94.1%; Pred. No. 9.8e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TCCATCTCATGTGAT 19
    |||||
Db 471 TCCATCTCATGTGAT 455

RESULT 42
US-10-424-599-103050
; Sequence 103050, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 103050
; LENGTH: 715
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MKT3847_64071C.1
US-10-424-599-103050

Query Match          77.0%; Score 15.4; DB 13; Length 715;
Best Local Similarity 94.1%; Pred. No. 1e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CCATCTCATGTGATG 20
    |||||
Db 471 TCCATCTCATGTGAT 455
```

Db 368 CCATGCTCATGTTGATG 384

## RESULT 43

US-10-424-599-23361/c  
; Sequence 23361, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223) B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 23361  
; LENGTH: 3767  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_121097C.1  
US-10-424-599-23361

Query Match 77.0%; Score 15.4; DB 13; Length 3767;  
Best Local Similarity 94.1%; Pred. No. 1.4e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGA 18  
Db 2838 GTCCATCTCATGTTGA 2822

## RESULT 44

US-10-087-192-1807/c  
; Sequence 1807, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1807  
; LENGTH: 119414  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(119414)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-1807

Query Match 77.0%; Score 15.4; DB 13; Length 119414;  
Best Local Similarity 94.1%; Pred. No. 2.5e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGA 18  
Db 11805 GTCCATCTCATGTTGA 11789

## RESULT 45

US-09-754-853A-2/c  
; Sequence 2, Application US/09754853A  
; Publication No. US20030005491A1  
; GENERAL INFORMATION:  
; APPLICANT: Hauge, Brian M.  
; APPLICANT: Parnell, Laurence D.  
; APPLICANT: Parsons, Jeremy D.  
; APPLICANT: Wang, Ming Li  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance  
; FILE REFERENCE: 38-10(15810) B  
; CURRENT APPLICATION NUMBER: US/09/754,853A  
; CURRENT FILING DATE: 2001-01-05  
; PRIOR APPLICATION NUMBER: US 60/174,880  
; PRIOR FILING DATE: 2000-01-07  
; NUMBER OF SEQ ID NOS: 1119  
; SEQ ID NO 2  
; LENGTH: 335913  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (45163)..(45314), (45450)..(45509), (46941)..(48763), (48975)..(49573)  
; OTHER INFORMATION: Clone ID: 240017\_region\_G3  
US-09-754-853A-2

Query Match 77.0%; Score 15.4; DB 10; Length 335913;  
Best Local Similarity 94.1%; Pred. No. 2.8e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATCTCATGTTGAT 19  
Db 24565 TCCAACTCATGTTGAT 24549

## RESULT 46

US-09-754-853A-3/c  
; Sequence 3, Application US/09754853A  
; Publication No. US20030005491A1  
; GENERAL INFORMATION:  
; APPLICANT: Hauge, Brian M.  
; APPLICANT: Parnell, Laurence D.  
; APPLICANT: Parsons, Jeremy D.  
; APPLICANT: Wang, Ming Li  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-10(15810) B  
; CURRENT APPLICATION NUMBER: US/09/754,853A  
; CURRENT FILING DATE: 2001-01-05  
; PRIOR APPLICATION NUMBER: US 60/174,880  
; PRIOR FILING DATE: 2000-01-07  
; NUMBER OF SEQ ID NOS: 1119  
; SEQ ID NO 3  
; LENGTH: 335913  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (46798)..(48763), (48975)..(49573)  
; OTHER INFORMATION: Clone ID: 240017\_region\_G3  
US-09-754-853A-3

Query Match 77.0%; Score 15.4; DB 10; Length 335913;  
Best Local Similarity 94.1%; Pred. No. 2.8e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATCTCATGTTGAT 19  
Db 24565 TCCAACTCATGTTGAT 24549

## RESULT 47

US-10-424-599-127062/c

```
; Sequence 127062, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 127062
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85743C.1
US-10-424-599-127062

Query Match          76.0%; Score 15.2; DB 13; Length 292;
Best Local Similarity 85.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GGTCATCTCATGTGTGATG 20
DB      37  GGTCATACCCATGGGTGGTG 18

RESULT 48
US-08-781-986A-1137
; Sequence 1137, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-1137

Query Match          76.0%; Score 15.2; DB 8; Length 330;
```

```
Best Local Similarity 85.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GGTCATCTCATGTGTGATG 20
DB      284  GGACCAAACTCATGTGTGATG 303

RESULT 49
US-10-329-624-1137
; Sequence 1137, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; APPLICANT: Gil H. Choi
; APPLICANT: Patrick S. Dillon
; APPLICANT: Craig A. Rosen
; APPLICANT: Steven C. Barash
; APPLICANT: Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1137:
US-10-329-624-1137

Query Match          76.0%; Score 15.2; DB 13; Length 330;
Best Local Similarity 85.0%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GGTCATCTCATGTGTGATG 20
DB      284  GGACCAAACTCATGTGTGATG 303

RESULT 50
US-10-424-599-130364
; Sequence 130364, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
```

APPLICANT: La Rosa Thomas J  
 APPLICANT: Kovalic David K  
 APPLICANT: Zhou Yihua  
 APPLICANT: Cao Yongwei  
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53223)B  
 CURRENT APPLICATION NUMBER: US/10/424,599  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 285684  
 SEQ ID NO 130364  
 LENGTH: 358  
 TYPE: DNA  
 ORGANISM: Glycine max  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT3847\_88727C.1  
 US-10-424-599-130364

Query Match	Score 15.2;	DB 13;	length 358;
Best Local Similarity	85.0%;	Pred. No. 1,1e+03;	
Matches 17;	Conservative 0;	Mismatches 3;	Indels 0;
			Gaps 0;
QY	1	GGTCATCTCATGTTGATG	20
Db	97	GGCTTAATCTTGTTGATG	116

Search completed: August 17, 2004, 16:56:43  
Job time : 396 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 13:20:27 ; Search time 1548 Seconds  
(without alignments)  
559.988 Million cell updates/sec

Title: US-09-825-489-3

Perfect score: 20

Sequence: 1 ggtccatctctgtgtgag 20

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%

Database :

Listing first 800 summaries

GenEmbl:

1: gb\_ba:\*

2: gb\_hlg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pac:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_scs:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pac:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_scs:\*

28: em\_sy:\*

29: em\_un:\*

30: em\_vl:\*

31: em\_hlg\_hum:\*

32: em\_hlg\_inv:\*

33: em\_hlg\_other:\*

34: em\_hlg\_mus:\*

35: em\_hlg\_pin:\*

36: em\_hlg\_rnd:\*

37: em\_hlg\_vrt:\*

38: em\_hlg\_vrt:\*

39: em\_hlg\_hum:\*

40: em\_hlg\_mus:\*

41: em\_hlg\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	AX282880	AX282880 Sequence
2	20	100.0	809	HSXPAC5	U10347 Human Xerod
3	20	100.0	822	AR183106	AR183106 Sequence
4	20	100.0	822	AR203243	AR203243 Sequence
5	20	100.0	822	AX277000	AX277000 Sequence
6	20	100.0	1377	AX282890	AX282890 Sequence
7	20	100.0	1377	HMXKPA6	D14533 Homo sapien
8	20	100.0	1439	BC014965	BC014965 Homo sapi
9	20	100.0	24994	AF503166	AF503166 Homo sapi
10	20	100.0	111345	AL445531	AL445531 Human DNA
11	20	100.0	187079	AL442130	AL442130 Homo sapi
12	18.4	92.0	1095	XELXPACB2	D31895 Xenopus lae
13	18.4	92.0	1277	XELXPACB	D31894 Xenopus lae
14	18.4	92.0	157653	AL732555	AL732555 Mouse DNA
15	18.4	92.0	199263	AC109694	AC109694 Rattus no
16	18.4	92.0	227524	AC126180	AC126180 Rattus no
17	17.4	87.0	158826	AP003607	AP003607 Oryza sat
18	17.4	87.0	197674	AP004367	AP004367 Oryza sat
19	17.4	87.0	198448	AC115118	AC115118 Mus muscu
20	17	85.0	189134	HDAC004682	AC004682 Homo sapi
21	17	85.0	190856	AC009160	AC009160 Homo sapi
22	17	85.0	253297	AC134076	AC134076 Rattus no
23	17	85.0	263730	AC094314	AC094314 Rattus no
24	16.8	84.0	294	MXKPA6	X74350 M. musculus
25	16.8	84.0	576	AF354052	AF354052 Mus muscu
26	16.8	84.0	938	MXKPA7	X74351 M. musculus
27	16.8	84.0	1029	AX566974	AX566974 Sequence
28	16.8	84.0	2138	AK115474	AK115474 Ciona int
29	16.8	84.0	11344	AE007333	AE007333 Streptoco
30	16.8	84.0	14273	AR218808	AR218808 Sequence
31	16.8	84.0	14273	BD003720	BD003720 Polynucle
32	16.8	84.0	90169	HS398A12	AL021332 Human DNA
33	16.8	84.0	110000	AC095009_1	Continuation (2 of
34	16.8	84.0	121113	BX537312	BX537312 Danio rer
35	16.8	84.0	121852	AP003965	AP003965 Homo sapi
36	16.8	84.0	127247	HS326133	Z82170 Human DNA
37	16.8	84.0	129300	AP003886	AP003886 Oryza sat
38	16.8	84.0	148416	AC002089	AC002089 Homo sapi
39	16.8	84.0	152635	AC011144	AC011144 Homo sapi
40	16.8	84.0	155646	AC146044	AC146044 Pan trogl
41	16.8	84.0	157200	AP005233	AP005233 Homo sapi
42	16.8	84.0	164572	OSJN00128	AL607001 Oryza sat
43	16.8	84.0	175936	SPNEU1908	AL499930 Streptoco
44	16.8	84.0	186184	AC139894	AC139894 Rattus no
45	16.8	84.0	209706	AC119611	AC119611 Rattus no
46	16.8	84.0	213381	AC129598	AC129598 Mus muscu
47	16.8	84.0	216911	AC107756	AC107756 Mus muscu
48	16.8	84.0	225182	AC107742	AC107742 Mus muscu
49	16.8	84.0	300340	AB016936	AB016936 Bacteroid
50	16.8	84.0	349980	AX571760	AX571760 Sequence
51	16.4	82.0	3573	LLJ000883	AJ000883 Lactococc
52	16.4	82.0	6765	AX586945	AX586945 Sequence
53	16.4	82.0	7568	AF092042	AF092042 Lactococc
54	16.4	82.0	10240	AB006382	AB006382 Lactococc
55	16.4	82.0	15246	AX586923	AX586923 Sequence
56	16.4	82.0	22868	AF536818	AF536818 Homo sapi
57	16.4	82.0	49321	AF275364S3	AF275366 Mus muscu
58	16.4	82.0	56758	AC101049	AC101049 Mus muscu
59	16.4	82.0	63489	AP002864	AP002864 Oryza sat
60	16.4	82.0	92487	AC068280	AC068280 Homo sapi
61	16.4	82.0	93388	AC105432	AC105432 Magnapor
62	16.4	82.0	93876	AL390240	AL390240 Human DNA
63	16.4	82.0	110000	AC114905_1	Continuation (2 of
64	16.4	82.0	115768	AP000712	AP000712 Homo sapi
65	16.4	82.0	137411	AC124678	AC124678 Mus muscu

C 66	16.4	82.0	137569	2	AC084093	AC084093 Homo sapi	C 139	15.8	79.0	1604	10	NMST6SIAD	X34000 M. musculus
C 67	16.4	82.0	140730	9	AP001960	AP001960 Homo sapi	C 140	15.8	79.0	1844	8	AK103896	AK103896 Oryza sat
C 68	16.4	82.0	141272	9	AP001976	AP001976 Homo sapi	C 141	15.8	79.0	2018	3	AY113567	AY113567 Drosophill
C 69	16.4	82.0	144383	8	AC092780	AC092780 Oryza sat	C 142	15.8	79.0	2282	8	BC000148	BC000148 Homo sapi
C 70	16.4	82.0	148501	9	BX284929	BX284929 Human DNA	C 143	15.8	79.0	2285	8	AK118438	AK118438 Arabidops
C 71	16.4	82.0	150394	2	BX005147	BX005147 Dantio rer	C 144	15.8	79.0	2302	9	BC003686	BC003686 Homo sapi
C 72	16.4	82.0	150948	8	AC091246	AC091246 Oryza sat	C 145	15.8	79.0	2324	9	BC022890	BC022890 Homo sapi
C 73	16.4	82.0	152858	8	AC093428	AC093428 Homo sapi	C 146	15.8	79.0	2712	7	BR034036	BR034036 Bacterioph
C 74	16.4	82.0	154937	2	AP002382	AP002382 Homo sapi	C 147	15.8	79.0	2963	3	AF435855	AF435855 Plasmodiu
C 75	16.4	82.0	159198	2	AP004866	AP004866 Oryza sat	C 148	15.8	79.0	3021	3	AF159136	AF159136 Dendrocto
C 76	16.4	82.0	159766	10	AC122292	AC122292 Mus muscu	C 149	15.8	79.0	4039	14	HSB182A	101099 Bovine herp
C 77	16.4	82.0	161338	10	AL645532	AL645532 Mouse DNA	C 150	15.8	79.0	5874	3	AF159137S2	AF159138 Dendrocto
C 78	16.4	82.0	164502	2	AC026960	AC026960 Homo sapi	C 151	15.8	79.0	10671	1	AE007842	AE007842 Clostrid
C 79	16.4	82.0	167084	8	AP002999	AP002999 Homo sapi	C 152	15.8	79.0	11034	14	AF104985	AF104985 Hixame rh
C 80	16.4	82.0	168552	8	AP004891	AP004891 Oryza sat	C 153	15.8	79.0	15033	2	AC012888	AC012888 Drosophill
C 81	16.4	82.0	168875	9	AC114794	AC114794 Homo sapi	C 154	15.8	79.0	18900	8	AY032722	AY032722 Homo sapi
C 82	16.4	82.0	169232	2	AC024299	AC024299 Homo sapi	C 155	15.8	79.0	20149	8	CNS0867M	AL731737 Oryza sat
C 83	16.4	82.0	170212	2	AC058822	AC058822 Homo sapi	C 156	15.8	79.0	24006	2	AC014743	AC014743 Drosophill
C 84	16.4	82.0	170878	2	BX469895	BX469895 Dantio rer	C 157	15.8	79.0	29545	9	AC112696	AC112696 Homo sapi
C 85	16.4	82.0	178367	2	AC015614	AC015614 Homo sapi	C 158	15.8	79.0	30583	2	AC017314	AC017314 Drosophill
C 86	16.4	82.0	181123	2	AC104064	AC104064 Homo sapi	C 159	15.8	79.0	35560	2	AC127494	AC127494 Homo sapi
C 87	16.4	82.0	182603	9	AF127019	AF127019 Homo sapi	C 160	15.8	79.0	50374	2	AC015170	AC015170 Drosophill
C 88	16.4	82.0	185768	10	AC102173	AC102173 Mus muscu	C 161	15.8	79.0	50607	2	AC090598	AC090598 Homo sapi
C 89	16.4	82.0	187131	2	AP001535	AP001535 Homo sapi	C 162	15.8	79.0	57287	2	AC008283	AC008283 Homo sapi
C 90	16.4	82.0	189927	10	AC120136	AC120136 Mus muscu	C 163	15.8	79.0	61173	1	AY372454	AY372454 Unculture
C 91	16.4	82.0	189949	2	AC110260	AC110260 Mus muscu	C 164	15.8	79.0	61532	2	AC144396	AC144396 Rattus no
C 92	16.4	82.0	192230	2	AC027065	AC027065 Homo sapi	C 165	15.8	79.0	61532	2	AC144396	AC144396 Rattus no
C 93	16.4	82.0	194518	9	AC113147	AC113147 Homo sapi	C 166	15.8	79.0	71396	2	AC100674	AC100674 Mus muscu
C 94	16.4	82.0	195410	9	AC096639	AC096639 Homo sapi	C 167	15.8	79.0	72387	2	AC005811	AC005811 Drosophill
C 95	16.4	82.0	195027	2	AC026611	AC026611 Homo sapi	C 168	15.8	79.0	73148	2	AC048349	AC048349 Homo sapi
C 96	16.4	82.0	204860	9	AC083902	AC083902 Homo sapi	C 169	15.8	79.0	73170	9	AL445068	AL445068 Human DNA
C 97	16.4	82.0	205594	9	AP002802	AP002802 Homo sapi	C 170	15.8	79.0	75499	9	HS0965G21	HS0965G21 Human DNA
C 98	16.4	82.0	207303	2	AL645562	AL645562 Homo sapi	C 171	15.8	79.0	86825	2	AC131082	AC131082 Mus muscu
C 99	16.4	82.0	211844	2	AC048370	AC048370 Homo sapi	C 172	15.8	79.0	90624	10	AL731671	AL731671 Mouse DNA
C 100	16.4	82.0	217293	9	AP001122	AP001122 Homo sapi	C 173	15.8	79.0	90882	2	AL139282	AL139282 Human DNA
C 101	16.4	82.0	226657	10	AL731843	AL731843 Mouse DNA	C 174	15.8	79.0	92059	2	AC128823_3	AC128823_3
C 102	16.4	82.0	234976	2	AC118698	AC118698 Mus muscu	C 175	15.8	79.0	92652	8	AC024261	AC024261 Arabidops
C 103	16.4	82.0	237845	2	AC112433	AC112433 Rattus no	C 176	15.8	79.0	92652	8	AL136363	AL136363 Human DNA
C 104	16.4	82.0	243479	2	BX537152	BX537152 Dantio rer	C 177	15.8	79.0	93296	9	CNS07Y00	AL732586 Human DNA
C 105	16.4	82.0	244441	2	AC098180	AC098180 Rattus no	C 178	15.8	79.0	99791	8	CNS07Y00	AL732586 Human DNA
C 106	16.4	82.0	249412	2	AC095097	AC095097 Rattus no	C 179	15.8	79.0	102092	2	AC136053	AC136053 Rattus no
C 107	16.4	82.0	254004	2	AC129469	AC129469 Rattus no	C 180	15.8	79.0	108873	14	AF136573	AF136573 Bovine he
C 108	16.4	82.0	257078	2	AC118300	AC118300 Rattus no	C 181	15.8	79.0	110000	2	AC098193_4	AC098193_4
C 109	16.4	82.0	257784	2	AC118627	AC118627 Mus muscu	C 182	15.8	79.0	110000	2	BX294172_1	BX294172_1
C 110	16.4	82.0	260480	2	AC116467	AC116467 Mus muscu	C 183	15.8	79.0	112405	9	AC072023	AC072023 Homo sapi
C 111	16.4	82.0	260657	2	AC094371	AC094371 Rattus no	C 184	15.8	79.0	112528	2	AC140068	AC140068 Medicago
C 112	16.4	82.0	300250	10	AF125314	AF125314 Mus muscu	C 185	15.8	79.0	119974	9	AC018763	AC018763 Homo sapi
C 113	16.4	82.0	51112	8	AF492377	AF492377 Aegilops	C 186	15.8	79.0	122079	9	AC009401	AC009401 Homo sapi
C 114	16.4	82.0	51112	9	AL445245	AL445245 Human DNA	C 187	15.8	79.0	122815	8	AC135563	AC135563 Oryza sat
C 115	16.4	82.0	81705	8	AP006384	AP006384 Lotus cor	C 188	15.8	79.0	125933	9	AL732586	AL732586 Human DNA
C 116	16.4	82.0	97864	8	AP006093	AP006093 Lotus cor	C 189	15.8	79.0	128758	1	AF279106	AF279106 Unculture
C 117	16.4	82.0	119684	2	AC008849	AC008849 Homo sapi	C 190	15.8	79.0	131133	8	AC146683	AC146683 Medicago
C 118	16.4	82.0	134319	9	AC113405	AC113405 Homo sapi	C 191	15.8	79.0	131441	8	CNS086CDA	AL9544871 Oryza sat
C 119	16.4	82.0	148415	2	AC113381	AC113381 Homo sapi	C 192	15.8	79.0	132221	2	AC102001	AC102001 Mus muscu
C 120	16.4	82.0	155895	2	AC102166	AC102166 Mus muscu	C 193	15.8	79.0	138129	2	AC011036	AC011036 Homo sapi
C 121	16.4	82.0	160703	9	AC022104	AC022104 Homo sapi	C 194	15.8	79.0	139788	2	AC024118	AC024118 Homo sapi
C 122	16.4	82.0	174101	2	AC119206	AC119206 Mus muscu	C 195	15.8	79.0	142347	10	AC125456	AC125456 Mus muscu
C 123	16.4	82.0	179424	2	AC027232	AC027232 Homo sapi	C 196	15.8	79.0	142712	9	AC112175	AC112175 Homo sapi
C 124	16.4	82.0	181792	9	AC098822	AC098822 Homo sapi	C 197	15.8	79.0	143451	2	AC021295	AC021295 Homo sapi
C 125	16.4	82.0	225164	2	AC104203	AC104203 Mus muscu	C 198	15.8	79.0	144357	2	AC144685	AC144685 Rattus no
C 126	16.4	82.0	228632	2	AC114202	AC114202 Rattus no	C 199	15.8	79.0	145247	10	AC138173	AC138173 Mus muscu
C 127	16.4	82.0	238919	2	AC139513	AC139513 Mus muscu	C 200	15.8	79.0	146128	9	AC026358	AC026358 Homo sapi
C 128	16.4	82.0	239893	2	AC107521	AC107521 Rattus no	C 201	15.8	79.0	147376	10	AL607039	AL607039 Mouse DNA
C 129	15.8	79.0	250	11	DM15812T	Z31891 D. melanoga	C 202	15.8	79.0	149546	9	AC087256	AC087256 Homo sapi
C 130	15.8	79.0	614	6	BD229387	BD229387 Human gen	C 203	15.8	79.0	149819	2	AC121091	AC121091 Mus muscu
C 131	15.8	79.0	666	6	AK380680	AK380680 Sequence	C 204	15.8	79.0	149960	5	AC144706	AC144706 Dantio rer
C 132	15.8	79.0	666	6	S8706889	S8706889 CD8 beta 1=	C 205	15.8	79.0	150106	2	AC144913	AC144913 Mus muscu
C 133	15.8	79.0	801	6	AR375131	AR375131 Sequence	C 206	15.8	79.0	150106	8	CNS086CE	AL732639 Oryza sat
C 134	15.8	79.0	936	6	AK060174	AK060174 Oryza sat	C 207	15.8	79.0	151547	8	AP003312	AP003312 Oryza sat
C 135	15.8	79.0	1176	6	AR376767	AR376767 Sequence	C 208	15.8	79.0	151862	10	AC102383	AC102383 Mus muscu
C 136	15.8	79.0	1490	8	AK069322	AK069322 Oryza sat	C 209	15.8	79.0	152220	8	OSJN00166	AL662967 Oryza sat
C 137	15.8	79.0	1698	8	AK102189	AK102189 Oryza sat	C 210	15.8	79.0	152313	9	AL591594	AL591594 Human DNA
C 138	15.8	79.0	1796	8	AK071708	AK071708 Oryza sat	C 211	15.8	79.0	153203	2	AL359812	AL359812 Homo sapi

212	15.8	79.0	153307	10	AL845441	285	15.8	79.0	191814	10	AC130541	AC130541 Mus muscu
C 213	15.8	79.0	154369	5	AC004803	C 286	15.8	79.0	191856	9	AL355980	AL355980 Human DNA
214	15.8	79.0	155713	5	BX000347	C 287	15.8	79.0	193391	2	AC026824	AC026824 Homo sapi
C 215	15.8	79.0	156225	2	AC115007	C 288	15.8	79.0	194224	5	AL954868	AL954868 Zebrafish
216	15.8	79.0	157310	2	RM187K23	C 289	15.8	79.0	194510	2	AC092956	AC092956 Homo sapi
217	15.8	79.0	157428	10	AL606832	C 290	15.8	79.0	195269	2	AC134458	AC134458 Homo sapi
218	15.8	79.0	158907	9	AC011890	C 291	15.8	79.0	195546	2	AC121701	AC121701 Rattus no
219	15.8	79.0	159235	9	AL356954	C 292	15.8	79.0	196350	2	BX248319	BX248319 Dantio rer
C 220	15.8	79.0	159305	9	AL355384	C 293	15.8	79.0	196817	2	AC019209	AC019209 Homo sapi
C 221	15.8	79.0	159919	9	AL359194	C 294	15.8	79.0	196983	2	BX323059	BX323059 Dantio rer
C 222	15.8	79.0	161067	9	AL353638	C 295	15.8	79.0	199289	10	AL772275	AL772275 Mouse DNA
C 223	15.8	79.0	161317	2	AL391873	C 296	15.8	79.0	199894	2	AC115053	AC115053 Mus muscu
C 224	15.8	79.0	161582	2	AC110650	C 297	15.8	79.0	200425	10	AC114549	AC114549 Mus muscu
C 225	15.8	79.0	163284	2	AL137000	C 298	15.8	79.0	200643	2	AC016418	AC016418 Homo sapi
C 226	15.8	79.0	163314	2	AL590557	C 299	15.8	79.0	202295	2	AC109142	AC109142 Mus muscu
C 227	15.8	79.0	163475	3	AC093099	C 300	15.8	79.0	203029	2	BX855594	BX855594 Mus muscu
C 228	15.8	79.0	164193	3	AC006494	C 301	15.8	79.0	203282	10	AC116674	AC116674 Mus muscu
C 229	15.8	79.0	165056	2	AC022412	C 302	15.8	79.0	204225	2	AC117423	AC117423 Homo sapi
C 230	15.8	79.0	165328	2	AC022185	C 303	15.8	79.0	204249	2	AC133268	AC133268 Rattus no
C 231	15.8	79.0	165589	2	AC011052	C 304	15.8	79.0	204445	2	AC126512	AC126512 Rattus no
C 232	15.8	79.0	165638	2	AC103514	C 305	15.8	79.0	205616	2	CNS05TDE	CNS05TDE
C 233	15.8	79.0	165964	2	AC128313	C 306	15.8	79.0	206431	2	AL935191	AL935191 Dantio rer
C 234	15.8	79.0	166185	5	BX649507	C 307	15.8	79.0	206438	5	AL935268	AL935268 Zebrafish
C 235	15.8	79.0	166652	2	BX537330	C 308	15.8	79.0	206784	9	HS9317	HS9317 Human DNA
C 236	15.8	79.0	167560	7	AY303349	C 309	15.8	79.0	206860	2	AC102884	AC102884 Mus muscu
C 237	15.8	79.0	168062	9	AL353768	C 310	15.8	79.0	208367	2	AC141868	AC141868 Mus muscu
C 238	15.8	79.0	168227	9	AC021868	C 311	15.8	79.0	208561	9	AC023206	AC023206 Homo sapi
C 239	15.8	79.0	168599	10	AL935336	C 312	15.8	79.0	209109	9	AC116025	AC116025 Homo sapi
C 240	15.8	79.0	168909	2	AC118862	C 313	15.8	79.0	209643	10	AC080018	AC080018 Mus muscu
C 241	15.8	79.0	169158	2	AC121376	C 314	15.8	79.0	210074	9	AC010837	AC010837 Homo sapi
C 242	15.8	79.0	171004	4	AC006286	C 315	15.8	79.0	210332	10	AL808131	AL808131 Mouse DNA
C 243	15.8	79.0	171209	4	AC124905	C 316	15.8	79.0	210537	2	AC109840	AC109840 Rattus no
C 244	15.8	79.0	171454	3	AC007137	C 317	15.8	79.0	211509	2	CNS08CBD	CNS08CBD
C 245	15.8	79.0	171520	2	BX470069	C 318	15.8	79.0	212475	2	AC093021	AC093021 Mus muscu
C 246	15.8	79.0	171710	2	AC013388	C 319	15.8	79.0	213045	2	AC120846	AC120846 Mus muscu
C 247	15.8	79.0	171868	3	AC010118	C 320	15.8	79.0	214505	2	AC099590	AC099590 Mus muscu
C 248	15.8	79.0	171970	2	AL590454	C 321	15.8	79.0	214838	9	AC012501	AC012501 Homo sapi
C 249	15.8	79.0	171974	9	AC092042	C 322	15.8	79.0	215294	2	AC117585	AC117585 Mus muscu
C 250	15.8	79.0	172123	9	AC060809	C 323	15.8	79.0	215615	2	AC098139	AC098139 Rattus no
C 251	15.8	79.0	172479	3	AC006590	C 324	15.8	79.0	219266	2	AC130080	AC130080 Rattus no
C 252	15.8	79.0	172695	9	AC068303	C 325	15.8	79.0	220367	2	AC087133	AC087133 Mus muscu
C 253	15.8	79.0	173535	9	AC142305	C 326	15.8	79.0	220935	2	AC111951	AC111951 Rattus no
C 254	15.8	79.0	173672	5	BX248086	C 327	15.8	79.0	224136	2	AC133260	AC133260 Rattus no
C 255	15.8	79.0	173844	9	AC078842	C 328	15.8	79.0	224417	2	AC095923	AC095923 Rattus no
C 256	15.8	79.0	173956	9	AC105383	C 329	15.8	79.0	224548	3	AE003739	AE003739 Drosophila
C 257	15.8	79.0	175222	2	AC132922	C 330	15.8	79.0	226579	10	AC133525	AC133525 Mus muscu
C 258	15.8	79.0	175238	2	AC101745	C 331	15.8	79.0	227155	10	AL732593	AL732593 Mouse DNA
C 259	15.8	79.0	176049	10	AL844198	C 332	15.8	79.0	230204	2	AC123996	AC123996 Mus muscu
C 260	15.8	79.0	176161	10	AC099578	C 333	15.8	79.0	231140	2	AC116231	AC116231 Rattus no
C 261	15.8	79.0	176987	2	AC103356	C 334	15.8	79.0	231165	2	AC120231	AC120231 Rattus no
C 262	15.8	79.0	177583	3	AC105264	C 335	15.8	79.0	233320	10	AC127314	AC127314 Mus muscu
C 263	15.8	79.0	178562	3	AC091954	C 336	15.8	79.0	233965	2	AC125739	AC125739 Rattus no
C 264	15.8	79.0	178931	3	AC008197	C 337	15.8	79.0	234429	9	AF311103	AF311103 Homo sapi
C 265	15.8	79.0	179199	2	AC120237	C 338	15.8	79.0	234758	2	AC130851	AC130851 Rattus no
C 266	15.8	79.0	180242	5	AL844185	C 339	15.8	79.0	234967	2	AC109380	AC109380 Rattus no
C 267	15.8	79.0	180872	2	AC145923	C 340	15.8	79.0	235008	2	AC119443	AC119443 Rattus no
C 268	15.8	79.0	181903	2	AC145923	C 341	15.8	79.0	235150	2	AC095650	AC095650 Rattus no
C 269	15.8	79.0	181935	9	AC018728	C 342	15.8	79.0	236832	2	AC119783	AC119783 Rattus no
C 270	15.8	79.0	182288	9	AC018362	C 343	15.8	79.0	237805	2	AC121047	AC121047 Rattus no
C 271	15.8	79.0	182452	9	AL161731	C 344	15.8	79.0	238355	2	AC126306	AC126306 Rattus no
C 272	15.8	79.0	186105	10	AC055777	C 345	15.8	79.0	239789	2	AC129342	AC129342 Rattus no
C 273	15.8	79.0	186564	10	AC118916	C 346	15.8	79.0	240747	2	AC105839	AC105839 Rattus no
C 274	15.8	79.0	187415	2	AC026352	C 347	15.8	79.0	241000	2	AC121369	AC121369 Rattus no
C 275	15.8	79.0	188078	2	AC008045	C 348	15.8	79.0	241834	2	AC095798	AC095798 Rattus no
C 276	15.8	79.0	188734	2	AC069443	C 349	15.8	79.0	241899	2	AC120228	AC120228 Rattus no
C 277	15.8	79.0	188734	2	AL357055	C 350	15.8	79.0	242037	2	AC114200	AC114200 Rattus no
C 278	15.8	79.0	188812	10	AC136091	C 351	15.8	79.0	242348	2	AC111361	AC111361 Rattus no
C 279	15.8	79.0	189650	9	AC023034	C 352	15.8	79.0	242386	2	AC098053	AC098053 Rattus no
C 280	15.8	79.0	190084	9	AC140950	C 353	15.8	79.0	242649	2	AC108648	AC108648 Rattus no
C 281	15.8	79.0	190692	9	AC021379	C 354	15.8	79.0	243563	2	AC098187	AC098187 Rattus no
C 282	15.8	79.0	190987	2	AC113883	C 355	15.8	79.0	244259	2	AC123226	AC123226 Rattus no
C 283	15.8	79.0	191239	2	AP000883	C 356	15.8	79.0	249148	2	AC106405	AC106405 Rattus no
C 284	15.8	79.0	191265	2	AC068996	C 357	15.8	79.0	249922	2	AC129001	AC129001 Rattus no

C 358	15.8	79.0	250087	2	AC096931	AC096931 Rattus no	431	15.4	77.0	144901	2	AC125276	AC125276 Mus muscu
359	15.8	79.0	250425	2	AC125667	AC125667 Rattus no	C 432	15.4	77.0	145536	2	AC105166	AC105166 Mus muscu
360	15.8	79.0	251173	2	AC115487	AC115487 Rattus no	C 433	15.4	77.0	149986	2	AC069256	AC069256 Homo sapi
361	15.8	79.0	251478	2	AC105849	AC105849 Rattus no	C 434	15.4	77.0	150076	9	CNS01DPI	AL121216 Human chr
362	15.8	79.0	252602	2	AC091704	AC091704 Mus muscu	C 435	15.4	77.0	153046	9	AL590009	AL590009 Human DNA
363	15.8	79.0	255185	2	AC132749	AC132749 Rattus no	436	15.4	77.0	154234	3	AC010034	AC010034 Drosophill
364	15.8	79.0	257867	2	AC005557	AC005557 Drosophill	C 437	15.4	77.0	157979	4	CNS05TD4	AL355922 Human chr
C 365	15.8	79.0	258416	2	AC098944	AC098944 Rattus no	438	15.4	77.0	158375	2	AC080039	AC080039 Homo sapi
C 366	15.8	79.0	258710	2	AC137281	AC137281 Rattus no	439	15.4	77.0	161602	2	AC022634	AC022634 Homo sapi
C 367	15.8	79.0	260027	3	AE003659	AE003659 Drosophill	440	15.4	77.0	162835	10	AC121866	AC121866 Mus muscu
C 368	15.8	79.0	273995	3	AE003669	AE003669 Drosophill	C 441	15.4	77.0	164932	10	AC121942	AC121942 Mus muscu
C 369	15.8	79.0	274027	2	AC134087	AC134087 Rattus no	442	15.4	77.0	166020	2	AC136149	AC136149 Oryza sat
C 370	15.8	79.0	284127	2	AC128719	AC128719 Rattus no	C 443	15.4	77.0	166138	2	AC084756	AC084756 Homo sapi
C 371	15.8	79.0	300134	1	AE001721	AE001721 Geobacter	C 444	15.4	77.0	166766	2	AC020685	AC020685 Homo sapi
C 372	15.8	79.0	303191	1	AE003472	AE003472 Drosophill	445	15.4	77.0	167676	2	AC119682	AC119682 Rattus no
373	15.8	79.0	310919	2	AC095296	AC095296 Rattus no	446	15.4	77.0	168862	2	AC092577	AC092577 Papio anu
374	15.8	79.0	315600	2	AL595304	AL595304 Homo sapi	447	15.4	77.0	168860	2	AC114315	AC114315 Homo sapi
375	15.4	77.0	633	6	AX142249	AX142249 Sequence	448	15.4	77.0	169757	5	AL627256	AL627256 Zebrafish
376	15.4	77.0	675	6	AX143961	AX143961 Sequence	C 449	15.4	77.0	170048	2	AC025240	AC025240 Homo sapi
377	15.4	77.0	754	10	BC048550	BC048550 Mus muscu	C 450	15.4	77.0	170466	3	AC010012	AC010012 Drosophill
C 378	15.4	77.0	895	8	AF526257	AF526257 Glycine m	C 451	15.4	77.0	170532	9	AL161728	AL161728 Human DNA
C 379	15.4	77.0	1420	8	AK026285	AK026285 Homo sapi	C 452	15.4	77.0	171347	9	AC099776	AC099776 Homo sapi
C 380	15.4	77.0	1686	8	HYUJ10841	HYUJ10841 Hordeum v	C 453	15.4	77.0	172295	10	AL845460	AL845460 Mouse DNA
C 381	15.4	77.0	1934	8	AT147012	AT147012 Medicago	C 454	15.4	77.0	173645	2	AC010770	AC010770 Homo sapi
C 382	15.4	77.0	2554	8	BD232696	BD232696 Encodemen	455	15.4	77.0	173769	9	AL696819	AL696819 Homo sapi
C 383	15.4	77.0	2876	8	AK065951	AK065951 Oryza sat	C 456	15.4	77.0	173962	2	AC024155	AC024155 Rattus no
384	15.4	77.0	2946	1	AF265698	AF265698 Staphyloc	C 457	15.4	77.0	174002	2	AC123406	AC123406 Rattus no
385	15.4	77.0	2946	6	AX145016	AX145016 Sequence	C 458	15.4	77.0	174393	3	AC011720	AC011720 Homo sapi
C 386	15.4	77.0	3021	1	AF269704	AF269704 Staphyloc	C 459	15.4	77.0	174949	2	AC111342	AC111342 Rattus no
C 387	15.4	77.0	3021	6	AX145022	AX145022 Sequence	C 460	15.4	77.0	175008	2	AC113300	AC113300 Mus muscu
388	15.4	77.0	3391	1	AF269717	AF269717 Staphyloc	C 461	15.4	77.0	175876	2	AC123336	AC123336 Rattus no
389	15.4	77.0	3391	6	AX145035	AX145035 Sequence	C 462	15.4	77.0	175946	9	AP002088	AP002088 Homo sapi
C 390	15.4	77.0	3725	10	BC042528	BC042528 Mus muscu	C 463	15.4	77.0	176021	5	AC026562	AC026562 Homo sapi
C 391	15.4	77.0	3768	8	AF195029	AF195029 Glycine m	C 464	15.4	77.0	176569	2	BX088540	BX088540 Zebrafish
C 392	15.4	77.0	11906	3	AF321003	AF321003 Homo sapi	C 465	15.4	77.0	176865	2	AL513306	AL513306 Homo sapi
393	15.4	77.0	14536	3	AE467260	AE467260 Ostrinia	C 466	15.4	77.0	176967	2	AC022742	AC022742 Homo sapi
C 394	15.4	77.0	39578	2	AC101097	AC101097 Mus muscu	C 467	15.4	77.0	177246	2	AC114308	AC114308 Homo sapi
C 395	15.4	77.0	49976	2	AC025670	AC025670 Homo sapi	C 468	15.4	77.0	177986	2	AC140745	AC140745 Rattus no
C 396	15.4	77.0	57144	2	AC100329	AC100329 Mus muscu	C 469	15.4	77.0	178316	9	AC022690	AC022690 Homo sapi
C 397	15.4	77.0	63753	2	AC090089	AC090089 Homo sapi	470	15.4	77.0	178364	10	AL645594	AL645594 Mouse DNA
C 398	15.4	77.0	64341	10	AL627323	AL627323 Mouse DNA	C 471	15.4	77.0	178899	9	AL589794	AL589794 Human DNA
399	15.4	77.0	65225	2	AC079109	AC079109 Homo sapi	C 472	15.4	77.0	179169	2	AC092759	AC092759 Papio anu
C 400	15.4	77.0	74282	8	AB025606	AB025606 Arabidops	C 473	15.4	77.0	179463	2	AL596453	AL596453 Homo sapi
C 401	15.4	77.0	75076	8	AC004948	AC004948 Homo sapi	474	15.4	77.0	179724	3	AP003094	AP003094 Homo sapi
C 402	15.4	77.0	79556	9	AC119567	AC119567 Homo sapi	C 475	15.4	77.0	183297	2	AC102312	AC102312 Mus muscu
C 403	15.4	77.0	80009	3	AF226688	AF226688 Bombyx mo	C 476	15.4	77.0	183547	2	AC009551	AC009551 Homo sapi
404	15.4	77.0	84653	2	BX663523	BX663523 Gallus ga	477	15.4	77.0	184194	9	AC008952	AC008952 Homo sapi
C 405	15.4	77.0	89666	2	AC022403	AC022403 Homo sapi	478	15.4	77.0	184688	9	AC092944	AC092944 Homo sapi
C 406	15.4	77.0	90015	9	AL513264	AL513264 Human DNA	479	15.4	77.0	185024	2	AC027649	AC027649 Mus muscu
C 407	15.4	77.0	90599	8	AC130275	AC130275 Medicago	C 480	15.4	77.0	185369	10	AL671860	AL671860 Mouse DNA
C 408	15.4	77.0	91073	9	AP001093	AP001093 Homo sapi	481	15.4	77.0	186233	9	AC092329	AC092329 Homo sapi
C 409	15.4	77.0	103177	5	BX005474	BX005474 Zebrafish	482	15.4	77.0	186253	9	AC020892	AC020892 Homo sapi
C 410	15.4	77.0	106095	2	BX663527	BX663527 Gallus ga	483	15.4	77.0	189317	2	AC024483	AC024483 Homo sapi
411	15.4	77.0	107093	2	AC097394	AC097394 Rattus no	484	15.4	77.0	189461	2	AC118247	AC118247 Mus muscu
C 412	15.4	77.0	110000	2	AC125752	AC125752 Drosophill	C 485	15.4	77.0	190718	8	CNS08C7T	AL171744 Oryza sat
C 413	15.4	77.0	110000	2	AC125752	AC125752 Drosophill	486	15.4	77.0	193931	2	AC116486	AC116486 Mus muscu
C 414	15.4	77.0	113024	2	AL445260	AL445260 Homo sapi	C 487	15.4	77.0	197288	2	AC115154	AC115154 Mus muscu
C 415	15.4	77.0	113311	2	AC019822	AC019822 Drosophill	C 488	15.4	77.0	197679	10	AC128663	AC128663 Mus muscu
C 416	15.4	77.0	114392	2	AC135880	AC135880 Rattus no	489	15.4	77.0	200500	9	AC093538	AC093538 Homo sapi
C 417	15.4	77.0	116130	9	AC026423	AC026423 Homo sapi	C 490	15.4	77.0	201106	2	AC110375	AC110375 Mus muscu
C 418	15.4	77.0	120088	9	AC004865	AC004865 Homo sapi	491	15.4	77.0	201964	10	MMNHG29N7	MMNHG29N7 Mus muscu
C 419	15.4	77.0	126128	10	CNS07YOX	AL713886 Mus muscu	492	15.4	77.0	201966	2	AC006289	AC006289 Homo sapi
C 420	15.4	77.0	128642	9	AC008838	AC008838 Homo sapi	493	15.4	77.0	202624	2	AC147075	AC147075 Pan trogl
C 421	15.4	77.0	128920	8	AC108871	AC108871 Oryza sat	494	15.4	77.0	202651	10	AC116887	AC116887 Mus muscu
C 422	15.4	77.0	129355	9	AC004853	AC004853 Homo sapi	495	15.4	77.0	203363	9	CNS01RMB	AL161668 Human chr
C 423	15.4	77.0	132348	10	CNS07YOZ	AL713882 Mus muscu	C 496	15.4	77.0	207901	10	AL590389	AL590389 Mouse DNA
C 424	15.4	77.0	134411	8	AL512662	AL512662 Human DNA	C 497	15.4	77.0	208734	10	AC102675	AC102675 Homo sapi
C 425	15.4	77.0	135378	8	CNS07YPP	AL113941 Oryza sat	498	15.4	77.0	211584	10	AC099101	AC099101 Rattus no
C 426	15.4	77.0	136037	8	AC004104	AC004104 Homo sapi	C 499	15.4	77.0	211931	10	AC125329	AC125329 Mus muscu
C 427	15.4	77.0	136777	2	AC146986	AC146986 Strongylo	C 500	15.4	77.0	211996	10	AC120837	AC120837 Mus muscu
C 428	15.4	77.0	139266	2	AC146775	AC146775 Medicago	501	15.4	77.0	213491	2	AC128375	AC128375 Rattus no
C 429	15.4	77.0	141428	9	AL139415	AL139415 Human DNA	C 502	15.4	77.0	213634	2	AC123360	AC123360 Mus muscu
430	15.4	77.0	143812	2	AC021847	AC021847 Homo sapi	503	15.4	77.0	214225	2	AC121430	AC121430 Rattus no



C 504	15.4	77.0	216568	2	ALJ71715	ALJ71715 Mus muscu	C 577	15.2	76.0	1035	6	AX349695	AX349695 Sequence
C 505	15.4	77.0	216957	2	AC119342	AC119342 Rattus no	C 578	15.2	76.0	1273	10	AB008911	AB008911 Mus muscu
C 506	15.4	77.0	219515	2	BX323558	BX323558 Danio rer	C 579	15.2	76.0	1287	6	AX122212	AX122212 Sequence
C 507	15.4	77.0	220807	2	AC139643	AC139643 Rattus no	C 580	15.2	76.0	1287	6	BD164329	BD164329 Novel pol
C 508	15.4	77.0	221952	2	AC098902	AC098902 Rattus no	C 581	15.2	76.0	1332	10	RNPNTS1	RNPNTS1
C 509	15.4	77.0	223102	2	AC094457	AC094457 Rattus no	C 582	15.2	76.0	1335	6	AX067039	AX067039 Sequence
C 510	15.4	77.0	224498	2	AC131973	AC131973 Mus muscu	C 583	15.2	76.0	1417	6	AX067037	AX067037 Sequence
C 511	15.4	77.0	225082	2	AC102806	AC102806 Mus muscu	C 584	15.2	76.0	1410	6	AX813887	AX813887 Sequence
C 512	15.4	77.0	226711	2	AC127440	AC127440 Rattus no	C 585	15.2	76.0	1419	6	AR388495	AR388495 Sequence
C 513	15.4	77.0	227029	2	AC125976	AC125976 Rattus no	C 586	15.2	76.0	1429	5	CHX8PAC4	CHX8PAC4
C 514	15.4	77.0	228330	10	AC095281	AC095281 Rattus no	C 587	15.2	76.0	1453	8	AF068934	AF068934
C 515	15.4	77.0	229784	10	ALB45171	ALB45171 Mouse DNA	C 588	15.2	76.0	1526	8	AF035700	AF035700
C 516	15.4	77.0	231175	2	AC094467	AC094467 Mouse DNA	C 589	15.2	76.0	1557	14	AF138298	AF138298
C 517	15.4	77.0	232225	10	AL662876	AL662876 Mouse DNA	C 590	15.2	76.0	1659	6	EO3827	EO3827
C 518	15.4	77.0	234477	2	AC129063	AC129063 Rattus no	C 591	15.2	76.0	1751	8	AF096247	AF096247
C 519	15.4	77.0	234528	2	AC096222	AC096222 Rattus no	C 592	15.2	76.0	1827	8	BT001108	BT001108
C 520	15.4	77.0	239308	2	AC107711	AC107711 Rattus no	C 593	15.2	76.0	1852	8	AY133701	AY133701
C 521	15.4	77.0	239559	2	AC125908	AC125908 Rattus no	C 594	15.2	76.0	2000	6	AX510247	AX510247
C 522	15.4	77.0	239570	2	AC137983	AC137983 Mus muscu	C 595	15.2	76.0	2000	6	AX652279	AX652279
C 523	15.4	77.0	239726	2	AC097728	AC097728 Rattus no	C 596	15.2	76.0	2161	1	AF411142	AF411142
C 524	15.4	77.0	241587	2	AC096197	AC096197 Rattus no	C 597	15.2	76.0	2180	9	AK126453	AK126453
C 525	15.4	77.0	242686	2	AC103159	AC103159 Rattus no	C 598	15.2	76.0	2317	8	AP170172	AP170172
C 526	15.4	77.0	247978	2	AC094450	AC094450 Rattus no	C 599	15.2	76.0	2319	10	MUSNFKBP4	MUSNFKBP4
C 527	15.4	77.0	248552	2	AC095542	AC095542 Rattus no	C 600	15.2	76.0	2368	6	BD218740	BD218740
C 528	15.4	77.0	253279	2	AC093477	AC093477 Mus muscu	C 601	15.2	76.0	2385	1	BRLCHOB	BRLCHOB
C 529	15.4	77.0	254295	2	BX663526	BX663526 Gallus ga	C 602	15.2	76.0	2385	6	EO3828	EO3828
C 530	15.4	77.0	254474	2	AC093478	AC093478 Mus muscu	C 603	15.2	76.0	2435	10	S66556	S66556 p70-ranact
C 531	15.4	77.0	256779	2	AC105578	AC105578 Rattus no	C 604	15.2	76.0	2483	10	S89033	S89033 I kappa B g
C 532	15.4	77.0	257239	2	AC115215	AC115215 Rattus no	C 605	15.2	76.0	2660	14	AY102920	AY102920 Vesticular
C 533	15.4	77.0	257405	2	AC135671	AC135671 Mus muscu	C 606	15.2	76.0	2799	10	RNPEANMT	RNPEANMT
C 534	15.4	77.0	257609	2	AC137477	AC137477 Rattus no	C 607	15.2	76.0	2810	9	BC021128	BC021128 Homo sapi
C 535	15.4	77.0	257959	2	AC094598	AC094598 Rattus no	C 608	15.2	76.0	2872	9	BC035072	BC035072 Homo sapi
C 536	15.4	77.0	259661	2	AC122570	AC122570 Rattus no	C 609	15.2	76.0	2915	6	AX695502	AX695502 Sequence
C 537	15.4	77.0	260625	2	AC096018	AC096018 Rattus no	C 610	15.2	76.0	2932	10	AB18195	AB18195 Mus muscu
C 538	15.4	77.0	261665	2	AC095565	AC095565 Rattus no	C 611	15.2	76.0	3324	6	AX675131	AX675131 Sequence
C 539	15.4	77.0	263387	2	AC112327	AC112327 Rattus no	C 612	15.2	76.0	3324	6	HSMB02237	HSMB02237
C 540	15.4	77.0	263388	2	AC094260	AC094260 Rattus no	C 613	15.2	76.0	3573	8	CHPSRPL2	CHPSRPL2
C 541	15.4	77.0	271135	2	AC103217	AC103217 Rattus no	C 614	15.2	76.0	3657	6	AX675130	AX675130
C 542	15.4	77.0	274751	2	BX649574	BX649574 Mus muscu	C 615	15.2	76.0	3663	6	AX675129	AX675129 Sequence
C 543	15.4	77.0	275392	2	AC095297	AC095297 Rattus no	C 616	15.2	76.0	3762	6	AX675147	AX675147 Sequence
C 544	15.4	77.0	278425	2	AC103101	AC103101 Rattus no	C 617	15.2	76.0	3768	6	AX675146	AX675146 Sequence
C 545	15.4	77.0	288495	2	AC095608	AC095608 Rattus no	C 618	15.2	76.0	3892	6	AX306113	AX306113 Sequence
C 546	15.4	77.0	291709	2	AC094701	AC094701 Rattus no	C 619	15.2	76.0	3892	6	AX695501	AX695501 Sequence
C 547	15.4	77.0	294316	2	AC131613	AC131613 Rattus no	C 620	15.2	76.0	3892	10	MUSPEPDS	MUSPEPDS
C 548	15.4	77.0	298959	2	AC128993	AC128993 Rattus no	C 621	15.2	76.0	3912	1	REO242746	REO242746
C 549	15.4	77.0	300892	1	AE016747	AE016747 Staphyloc	C 622	15.2	76.0	4114	10	AF110477	AF110477 Rattus no
C 550	15.4	77.0	302835	1	AE012555	AE012555 Xylella f	C 623	15.2	76.0	4270	3	AF070482	AF070482 Lytechinu
C 551	15.4	77.0	311857	2	BX000982	BX000982 Mus muscu	C 624	15.2	76.0	4304	10	AF110478	AF110478 Rattus no
C 552	15.4	77.0	319072	3	AE003474	AE003474 Drosophi1	C 625	15.2	76.0	4330	9	AB052103	AB052103 Homo sapi
C 553	15.4	77.0	335913	6	AX196295	AX196295 Sequence	C 626	15.2	76.0	4591	14	ASU13763	ASU13763 African swi
C 554	15.4	77.0	335913	6	AX196295	AX196295 Sequence	C 627	15.2	76.0	4884	6	AR319928	AR319928 Sequence
C 555	15.2	76.0	170	14	VSVD15A	M10402 VSV defecti	C 628	15.2	76.0	5321	8	AF116598	AF116598 Vitis vin
C 556	15.2	76.0	177	14	RHVS0V6	X00464 5'-region o	C 629	15.2	76.0	5327	6	AX100077	AX100077
C 557	15.2	76.0	259	11	G63946	G63946 UMBT1.003 Bo	C 630	15.2	76.0	5455	2	AC069012	AC069012 Mus muscu
C 558	15.2	76.0	259	11	KLAJ9990	AXJ229990 Kluverver	C 631	15.2	76.0	5661	9	AF113514	AF113514
C 559	15.2	76.0	308	6	AX920060	AX920060 Sequence	C 632	15.2	76.0	5881	6	AX807249	AX807249 Sequence
C 560	15.2	76.0	308	6	BD055583	BD055583 Sequence	C 633	15.2	76.0	5881	6	AX807415	AX807415 Sequence
C 561	15.2	76.0	330	6	AR355019	AR355019 Sequence	C 634	15.2	76.0	5909	9	AB061870	AB061870 Macaca fa
C 562	15.2	76.0	366	14	VSVDGL	K03386 Vesticular s	C 635	15.2	76.0	5988	9	AF119230	AF119230 Homo sapi
C 563	15.2	76.0	422	11	G24310	G24310 human STS W	C 636	15.2	76.0	6222	6	AX695649	AX695649 Sequence
C 564	15.2	76.0	441	14	VSVDLITG	K01682 Vesticular s	C 637	15.2	76.0	6330	6	AX100076	AX100076 Sequence
C 565	15.2	76.0	443	6	AX675141	AX675141 Sequence	C 638	15.2	76.0	6380	6	AX100079	AX100079 Sequence
C 566	15.2	76.0	450	8	AJ587452	AJ587452 Arabidops	C 639	15.2	76.0	6380	14	RHVS0POL	RHVS0POL
C 567	15.2	76.0	472	6	AX675139	AX675139 Sequence	C 640	15.2	76.0	6537	6	AX695648	AX695648 Sequence
C 568	15.2	76.0	499	6	AX675137	AX675137 Sequence	C 641	15.2	76.0	6537	9	AF119231	AF119231 Homo sapi
C 569	15.2	76.0	548	6	AX675137	AX675137 Sequence	C 642	15.2	76.0	8293	9	AB002381	AB002381 Human mRN
C 570	15.2	76.0	581	6	AX675136	AX675136 Sequence	C 643	15.2	76.0	8592	9	AF217500	AF217500 Homo sapi
C 571	15.2	76.0	608	11	BV017953	BV017953 S209P6355	C 644	15.2	76.0	8592	9	DMU77947	DMU77947
C 572	15.2	76.0	608	11	BV043721	BV043721 S212P6535	C 645	15.2	76.0	8401	3	AY095952	AY095952 Bortetell
C 573	15.2	76.0	632	11	BV070174	BV070174 S212P6118	C 646	15.2	76.0	8877	1	AY095952	AY095952 Bortetell
C 574	15.2	76.0	677	14	VSVDHDLT	K01640 Vesticular s	C 647	15.2	76.0	9196	5	GGTGRB2E5	GGTGRB2E5
C 575	15.2	76.0	708	6	AX675135	AX675135 Sequence	C 648	15.2	76.0	10618	2	AC015408	AC015408 Drosophi1
C 576	15.2	76.0	859	11	BV065920	BV065920 S212P6911	C 649	15.2	76.0	11161	14	VSVCG	J02428 Vesticular s

650	15.2	76.0	11162	14	AF473864	AF473864 Vesicular	C 723	15.2	76.0	11698	10	AL731687	AL731687 Mouse DNA
651	15.2	76.0	11167	1	AE001599	AE001599 Chlamydia	C 724	15.2	76.0	116951	8	AC114788	AC114788 Homo sapi
652	15.2	76.0	11168	1	AE002222	AE002222 Chlamydia	C 725	15.2	76.0	117226	8	AP006626	AP006626 Oryza sat
653	15.2	76.0	11688	1	AE009887	AE009887 Pyrobacul	C 726	15.2	76.0	118235	10	AC003060	AC003060 Mus muscu
654	15.2	76.0	12132	1	AE008393	AE008393 Streptoco	C 727	15.2	76.0	118313	9	AC003991	AC003991 Homo BAC
655	15.2	76.0	13978	8	AB013613	AB013613 Oryza sat	C 728	15.2	76.0	119666	9	HSD51B2	HSD51B2 Homo sapi
656	15.2	76.0	14179	1	AE007053	AE007053 Mycobacte	C 729	15.2	76.0	121961	8	NCB23H20	NCB23H20 Neurospor
657	15.2	76.0	14311	6	AR123015	AR123015 Sequence	C 730	15.2	76.0	122151	10	AL845310	AL845310 Mouse DNA
658	15.2	76.0	14311	6	AR123015	AR123015 Sequence	C 731	15.2	76.0	122356	10	AC093087	AC093087 Homo sapi
659	15.2	76.0	16332	2	AC023615	AC023615 Mus muscu	C 732	15.2	76.0	123039	8	AC123518	AC123518 Oryza sat
660	15.2	76.0	19300	14	AF327841	AF327841 African B	C 733	15.2	76.0	125008	10	AB005546	AB005546 Oryza sat
661	15.2	76.0	20868	2	AC017847	AC017847 Drosophil	C 734	15.2	76.0	127220	10	BX284649	BX284649 Mouse DNA
662	15.2	76.0	23009	1	AE008763	AE008763 Salmonell	C 735	15.2	76.0	127220	10	BX284649	BX284649 Mouse DNA
663	15.2	76.0	25967	2	AC100971	AC100971 Mus muscu	C 736	15.2	76.0	127334	9	AC109579	AC109579 Homo sapi
664	15.2	76.0	30875	2	AC0220A6	AC0220A6 Human DNA s	C 737	15.2	76.0	127790	9	AC008885	AC008885 Homo sapi
665	15.2	76.0	34051	9	AC104668	AC104668 Homo sapi	C 738	15.2	76.0	128580	2	AP005466	AP005466 Oryza sat
666	15.2	76.0	34713	2	BX539329	BX539329 Homo sapi	C 739	15.2	76.0	128683	9	AC027329	AC027329 Homo sapi
667	15.2	76.0	36411	1	U000022	U000022 Mycobacteri	C 740	15.2	76.0	128757	9	ALJ35937	ALJ35937 Human DNA
668	15.2	76.0	36412	6	AR345364	AR345364 Sequence	C 741	15.2	76.0	128856	2	AC126779	AC126779 Medicago
669	15.2	76.0	38322	14	AF327839	AF327839 African B	C 742	15.2	76.0	131098	2	AC146865	AC146865 Medicago
670	15.2	76.0	39089	3	CBRG33D04	CBRG33D04 Caenorhab	C 743	15.2	76.0	131427	14	OPU75930	OPU75930 Oryza p
671	15.2	76.0	40520	9	AC011513	AC011513 Homo sapi	C 744	15.2	76.0	131995	9	AC008496	AC008496 Homo sapi
672	15.2	76.0	40866	9	ALJ59033	ALJ59033 Human DNA	C 745	15.2	76.0	135434	9	ALJ354674	ALJ354674 Human DNA
673	15.2	76.0	43593	9	AC022516	AC022516 Homo sapi	C 746	15.2	76.0	137347	9	ALJ354674	ALJ354674 Human DNA
674	15.2	76.0	44806	9	AC069162	AC069162 Homo sapi	C 747	15.2	76.0	137908	10	AL607024	AL607024 Mouse DNA
675	15.2	76.0	45289	3	CBRG39M24	CBRG39M24 Caenorhab	C 748	15.2	76.0	138070	2	AC109783	AC109783 Mus muscu
676	15.2	76.0	47338	10	AF326469	AF326469 Mus muscu	C 749	15.2	76.0	138621	2	AP002028	AP002028 Homo sapi
677	15.2	76.0	47506	2	AC018082	AC018082 Drosophil	C 750	15.2	76.0	138982	2	AC104550	AC104550 Mus muscu
678	15.2	76.0	49356	1	AC025950	AC025950 Staphyloc	C 751	15.2	76.0	139443	2	AC066602	AC066602 Homo sapi
679	15.2	76.0	50152	1	AC025948	AC025948 Staphyloc	C 752	15.2	76.0	139636	10	AC124376	AC124376 Mus muscu
680	15.2	76.0	51410	9	AC092604	AC092604 Homo sapi	C 753	15.2	76.0	141428	2	AP005558	AP005558 Oryza sat
681	15.2	76.0	53826	2	AC145962	AC145962 Gallus ga	C 754	15.2	76.0	141993	2	AC121253	AC121253 Homo sapi
682	15.2	76.0	54157	8	T71414	AC005332 Arabidops	C 755	15.2	76.0	142291	2	AC025364	AC025364 Homo sapi
683	15.2	76.0	60801	9	AC006551	AC006551 Homo sapi	C 756	15.2	76.0	142717	2	AC079578	AC079578 Mus muscu
684	15.2	76.0	61579	9	ALJ36258	ALJ36258 Human DNA	C 757	15.2	76.0	143384	2	AC129050	AC129050 Rattus no
685	15.2	76.0	62794	8	AP003957	AP003957 Oryza sat	C 758	15.2	76.0	143588	10	BX005168	BX005168 Mouse DNA
686	15.2	76.0	63501	2	AC134974	AC134974 Homo sapi	C 759	15.2	76.0	143679	9	AC106793	AC106793 Homo sapi
687	15.2	76.0	63604	8	AP000381	AP000381 Arabidops	C 760	15.2	76.0	144883	2	CNS01DSL	CNS01DSL Human chr
688	15.2	76.0	69019	2	AC005117	AC005117 Drosophil	C 761	15.2	76.0	144883	2	AC091701	AC091701 Trypanoso
689	15.2	76.0	70872	2	AC114997	AC114997 Mus muscu	C 762	15.2	76.0	144503	2	AC100744	AC100744 Mus muscu
690	15.2	76.0	71044	2	AC024368	AC024368 Homo sapi	C 763	15.2	76.0	144651	8	AF229187	AF229187 Oryza sat
691	15.2	76.0	74512	2	AC100521	AC100521 Mus muscu	C 764	15.2	76.0	145370	2	AC116245	AC116245 Rattus no
692	15.2	76.0	74675	8	AP004787	AP004787 Oryza sat	C 765	15.2	76.0	146040	10	AC012399	AC012399 Homo sapi
693	15.2	76.0	78507	2	AC019704	AC019704 Drosophil	C 766	15.2	76.0	146554	2	BX855605	BX855605 Danio rer
694	15.2	76.0	79877	2	AC131495	AC131495 Lytechinu	C 767	15.2	76.0	146699	2	AC112210	AC112210 Homo sapi
695	15.2	76.0	82059	9	AC130896	AC130896 Homo sapi	C 768	15.2	76.0	146871	2	BX842689	BX842689 Danio rer
696	15.2	76.0	82080	10	AL844553	AL844553 Mouse DNA	C 769	15.2	76.0	147746	2	AC115874	AC115874 Homo sapi
697	15.2	76.0	82933	6	AX704274	AX704274 Sequence	C 770	15.2	76.0	147946	2	AC091432	AC091432 Homo sapi
698	15.2	76.0	85339	9	ALJ39086	ALJ39086 Human DNA	C 771	15.2	76.0	148320	10	AL772207	AL772207 Mouse DNA
699	15.2	76.0	85641	9	AC005217	AC005217 Homo sapi	C 772	15.2	76.0	148325	5	BX005124	BX005124 Zebrafish
700	15.2	76.0	86146	2	AC023805	AC023805 Mus muscu	C 773	15.2	76.0	149748	2	AC136287	AC136287 Mus muscu
701	15.2	76.0	88401	8	P23M19	AC007455 Arabidops	C 774	15.2	76.0	149844	2	AC144682	AC144682 Rattus no
702	15.2	76.0	91825	2	AC019854	AC019854 Drosophil	C 775	15.2	76.0	149995	2	AC016722	AC016722 Homo sapi
703	15.2	76.0	92562	10	AL732485	AL732485 Mouse DNA	C 776	15.2	76.0	150292	10	AC112342	AC112342 Rattus no
704	15.2	76.0	96079	9	HS717L17	AL0121706 Human DNA	C 777	15.2	76.0	150755	2	AC021258	AC021258 Homo sapi
705	15.2	76.0	96598	6	AX695500	AX695500 Sequence	C 778	15.2	76.0	151073	2	AC134624	AC134624 Oryza sat
706	15.2	76.0	97992	9	ALJ596118	ALJ596118 Human DNA	C 779	15.2	76.0	151289	2	AC102184	AC102184 Homo sapi
707	15.2	76.0	106123	2	AC073763	AC073763 Human DNA	C 780	15.2	76.0	151559	9	ALJ390294	ALJ390294 Human DNA
708	15.2	76.0	109770	2	AC095248	AC095248 Rattus no	C 781	15.2	76.0	152152	2	AC113505	AC113505 Mus muscu
709	15.2	76.0	110000	2	AC096079	AC096079 Rattus no	C 782	15.2	76.0	152304	2	AC023032	AC023032 Homo sapi
710	15.2	76.0	110000	2	AC098640	AC098640 Rattus no	C 783	15.2	76.0	152820	2	AC010210	AC010210 Homo sapi
711	15.2	76.0	110000	2	AC103355	AC103355 Rattus no	C 784	15.2	76.0	153080	2	AC026229	AC026229 Homo sapi
712	15.2	76.0	110000	2	AC110355	AC110355 Rattus no	C 785	15.2	76.0	153520	10	AC122520	AC122520 Mus muscu
713	15.2	76.0	110000	2	AC126130	AC126130 Rattus no	C 786	15.2	76.0	153994	8	AC138023	AC138023 Homo sapi
714	15.2	76.0	110000	2	BM572642	BM572642 Rattus no	C 787	15.2	76.0	156534	8	AP000391	AP000391 Oryza sat
715	15.2	76.0	110000	2	LMF52642	LMF52642 Rattus no	C 788	15.2	76.0	156547	10	AC117262	AC117262 Homo sapi
716	15.2	76.0	110000	6	AR310754	AR310754 Sequence	C 789	15.2	76.0	157394	2	AP004874	AP004874 Oryza sat
717	15.2	76.0	110102	8	ATP28172	ATP28172 Arabidops	C 790	15.2	76.0	157994	2	AL021770	AL021770 Arabidops
718	15.2	76.0	111354	8	AP003806	AP003806 Oryza sat	C 791	15.2	76.0	158276	9	AL160267	AL160267 Human DNA
719	15.2	76.0	111568	2	AP003908	AP003908 Homo sapi	C 792	15.2	76.0	158315	9	AC114478	AC114478 Homo sapi
720	15.2	76.0	112855	2	AP003608	AP003608 Oryza sat	C 793	15.2	76.0	158406	9	AC114297	AC114297 Homo sapi
721	15.2	76.0	115289	9	AC011490	AC011490 Homo sapi	C 794	15.2	76.0	159070	2	AC128114	AC128114 Rattus no
722	15.2	76.0	116653	2	AP005574	AP005574 Oryza sat	C 795	15.2	76.0	159104	9	AC108925	AC108925 Homo sapi

796 15.2 76.0 159468 2 AC023314 Homo sapi  
c 797 15.2 76.0 159519 2 AC018511 Homo sapi  
c 798 15.2 76.0 159617 9 AC022018 Homo sapi  
799 15.2 76.0 159832 2 AC104199 Mus muscu  
c 800 15.2 76.0 160039 9 AC104851 Homo sapi

## ALIGNMENTS

RESULT 1  
AX282880 20 bp DNA linear PAT 02-NOV-2001  
LOCUS Sequence 3 from Patent WO0174346.  
ACCESSION AX282880  
VERSION AX282880.1 GI:16609856  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Agrawal, S. and Kandimala, E. R.  
TITLE Sensitization of cells to cytotoxic agents using oligonucleotides directed to nucleotide excision repair or transcription coupled repair genes  
JOURNAL Patent: WO 0174346-A 3 11-OCT-2001;  
FEATURES  
source HYBRIDON, INC. (US)  
location/Qualifiers  
1..20

ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCCATCTCATGTTGATG 20  
Db 1 GGTCCATCTCATGTTGATG 20

RESULT 2  
HSXPAC5 809 bp DNA linear PRI 27-MAR-1996  
LOCUS Human Xeroderma Pigmentosum group A correcting (XPAC) gene, exon 6  
DEFINITION  
ACCESSION U10347  
VERSION U10347.1 GI:501113  
KEYWORDS  
SEGMENT 5 of 5  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS Topping, R. S., Myrand, S. P., Williams, B. L., Albert, J. C. and States, J. C.  
TITLE Characterization of the human XPA promoter  
JOURNAL Gene 166 (2), 341-342 (1995)  
MEDLINE 96125219  
PUBMED 8543191  
REFERENCE 2 (bases 1 to 809)  
AUTHORS States, J. C.  
TITLE Direct Submision  
JOURNAL Submitted (03-JUN-1994) J. Christopher States, Wayne State University, Center For Molecular Biology, 2727 Second Ave., Detroit, MI 48201, USA  
FEATURES  
source Location/Qualifiers  
1..809  
/organism="Homo sapiens"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/clone="Lambda XPAC 1"  
/sex="male"  
/tissue\_type="placental"  
/clone\_lib="Lambda FIX II Genomic DNA library from Stratagene"  
join(U10343.1:88..298,U10344.1:1..306,U10345.1:1..366,U10346.1:1..318,1..754)  
/gene="XPAC"  
join(U10343.1:<88..198,U10344.1:101..206,U10345.1:101..266,U10346.1:101..218,77..225)  
/gene="XPAC"  
/codon\_start=3  
/protein\_id="AA92883.1"  
/db\_xref="GI:501115"  
/translation="MANVKAAPKIIDNGGFTLEEEEBEOKIGVYHQRPWFYD VICECGEPMDSYMANHFDLPICDNCADADKHLITTEAKOBYLTKDCEKREP PLKFTVKKNPHHSQWDMKLYLKQIVRSLEWGSQALBEAKERQENRKMKQK FDKKVKELRRRAVRSVWKRRTIVHQHEYPEENLEDMDYRRTCTWCGHLYEKW"  
77..754  
/gene="XPAC"  
/number=6  
267..272  
/gene="XPAC"  
729..734  
/gene="XPAC"

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 809;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCCATCTCATGTTGATG 20  
Db 146 GGTCCATCTCATGTTGATG 127

RESULT 3  
AR183106/c 822 bp DNA linear PAT 20-APR-2002  
LOCUS AR183106  
DEFINITION Sequence 8 from patent US 6340566.  
ACCESSION AR183106  
VERSION AR183106.1 GI:20226699  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 822)  
TITLE McCutchen-Maloney, S. L.  
JOURNAL Detection and quantitation of single nucleotide polymorphisms, DNA sequence variations, DNA mutations, DNA damage and DNA mismatches  
PUBMED Patent: US 6340566-A 8 22-JAN-2002;  
FEATURES  
source Location/Qualifiers  
1..822  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 822;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCCATCTCATGTTGATG 20  
Db 743 GGTCCATCTCATGTTGATG 724

RESULT 4  
AR203243 822 bp DNA linear PAT 20-JUN-2002  
LOCUS AR203243/c  
DEFINITION Sequence 8 from patent US 6365355.  
ACCESSION AR203243

VERSION AR203243.1 GI:21499581  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 822)  
AUTHORS McCutchen-Maloney,S.L.  
TITLE Chimeric proteins for detection and quantitation of DNA mutations,  
DNA sequence variations, DNA damage and DNA mismatches  
JOURNAL Patent: US 6365355-A 8 02-APR-2002;  
FEATURES  
source Location/Qualifiers  
1..822  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 822;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
|||||  
743 GGTCCTACTCATGTTGATG 724

RESULT 5  
AX277000/c 822 bp DNA linear PAT 29-OCT-2001  
LOCUS AX277000/c  
DEFINITION Sequence 8 from Patent WO0173079.  
ACCESSION AX277000  
VERSION AX277000.1 GI:16548670  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS McCutchen-Maloney,S.L.  
TITLE Chimeric proteins for detection and quantitation of dna mutations,  
dna sequence variations, dna damage and dna mismatches  
JOURNAL Patent: WO 0173079-A 8 04-OCT-2001;  
The Regents of The University of California (US)  
FEATURES  
source Location/Qualifiers  
1..822  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 822;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
|||||  
743 GGTCCTACTCATGTTGATG 724

RESULT 6  
AX282890 1377 bp DNA linear PAT 02-NOV-2001  
LOCUS AX282890/c  
DEFINITION Sequence 13 from Patent WO0174346.  
ACCESSION AX282890  
VERSION AX282890.1 GI:16609866  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Agrawal,S. and Kandimala,E.R.  
TITLE Sensitization of cells to cytotoxic agents using oligonucleotides  
directed to nucleotide excision repair or transcription coupled

JOURNAL repair genes  
Patent: WO 0174346-A 13 11-OCT-2001;  
HYBRIDON, INC. (US)  
FEATURES  
source Location/Qualifiers  
1..1377  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 1377;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTTGATG 20  
|||||  
769 GGTCATCTCATGTTGATG 750

RESULT 7  
HMMXPAC/c 1377 bp mRNA linear PRI 11-FEB-2003  
LOCUS HMMXPAC/c  
DEFINITION Homo sapiens mRNA for XPAC protein, complete cds.  
ACCESSION D14533  
VERSION D14533.1 GI:286028  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Tanaka,K., Miura,N., Satokata,I., Miyamoto,I., Yoshida,M.C.,  
Tanaka,K., Kondo,S., Yasui,A., Okayama,H. and Okada,Y.  
TITLE Analysis of a human DNA excision repair gene involved in group A  
xeroderma pigmentosum and containing a zinc-finger domain  
JOURNAL Nature 348 (6296), 73-76 (1990)  
MEDLINE 91043046  
PUBMED 2234061  
REFERENCE 2  
AUTHORS Satokata,I., Iwai,K., Matasuda,T., Okada,Y. and Tanaka,K.  
TITLE Genomic characterization of the human DNA excision  
repair-controlling gene XPAC  
JOURNAL Gene 136 (1-2), 345-348 (1993)  
MEDLINE 94124028  
PUBMED 8294029  
REFERENCE 3 (bases 1 to 1377)  
AUTHORS Tanaka,K.  
TITLE Direct Submission  
JOURNAL Submitted (26-FEB-1993) Kiyoji Tanaka, Osaka University, Inst. for  
Molecular and Cellular Biology; 1-3 Yamadaoka, Suita, Osaka 565,  
Japan (Tel:81-6-6877-5238, Fax:81-6-6877-9136)  
FEATURES  
source Location/Qualifiers  
1..1377  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/clone\_lib="pcd2Basinger"  
/note="clones pcd2h19 and pcd2h29-human primer y  
fibroblast"  
1..1377  
/gene="XPAC"  
27..848  
/gene="XPAC"  
/codon\_start=1  
/product="XPAC protein"  
/protein\_id="BAA03403.1"  
/db\_xref="GI:286029"  
translation="MAADGALPEALALPEALPEASVRSASIRKROALMLRQARLA  
APYSATATAATGMAVAVKAPKTIIDGGGPILEEEBEEQKIGKVVHPQPVMEFDY  
VTEBCGKEPMDSITVMNHFDPPTDCNCRDADDKRLITKTBAKQBYLLKDDLEKRP  
PLKFTVKNPSPHSGWDMKLYLKQIVRSLELVGSGALAEAKAEKROENRKAKQAK  
FDKKVKELRRAVRSVWKRRTI VHQHEYGPENLEDMDYRRTCTMCGHELYEYEM"

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1377;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
 |||||  
 DB 769 GGTCCATCTCATGTTGATG 750

RESULT 8  
BC014965/c

LOCUS BC014965 1439 bp mRNA linear PRI 07-OCT-2003  
 DEFINITION Homo sapiens xeroderma pigmentosum, complementation group A, mRNA  
 (CDNA clone MGC:22059 IMAGE:4878739), complete cds.

ACCESSION BC014965  
 VERSION BC014965.1 GI:15929009  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1439)  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Mak, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldi, M.F., Caravanti, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Uedini, T.B., Toshiyuki, S.,  
 Carinci, P., Prange, C., Raha, S., Loggellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Meyer, R.M.,  
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,  
 Scherich, A., Schein, J.E., Jones, S.J., and Marra, M.A.

## TITLE

human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE 22388257  
 PUBMED 12477932  
 REFERENCE 2 (bases 1 to 1439)  
 Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK COMMENT  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabds-remail.nih.gov](mailto:cgabds-remail.nih.gov)  
 Tissue Procurement: Louis Staudt  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
 info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
 Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reia Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Mathewson, Candice Mcleavy, Steven  
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline  
 Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott,  
 Michael Thorne, Miranada Tsai, Natasia van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRL Plate: 34 Row: n Column: 16  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 4507936.  
 Location/Qualifiers  
 1..1439

## FEATURES

## source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:22059 IMAGE:4878739"  
 /tissue\_type="Primary B-Cells from Tonsils"  
 /clone\_id="NTH MGC 48"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 1..1439  
 /gene="XPA"  
 /note="synonyms: XPI, XPAC"  
 /db\_xref="LocusID:7507"  
 /db\_xref="MIM:278700"  
 66..887  
 /codon\_start=1  
 /product="xeroderma pigmentosum, complementation group A"  
 /protein\_id="AAH14965.1"  
 /db\_xref="GI:15929010"  
 /db\_xref="LocusID:7507"  
 /translation="MAADGALPEAALPEAPALPASVRSASIERKROALMLRQRLA  
 ARPSATAAATGCAKAKAPKILIDCGGILIEEPEERKIKVHOPVWEPDY  
 VICEGGERPDSTIMNFDPPTDNCADADKHLITKAKOYLKDDLEREP  
 PLKTVKKNPHRSQGMKLYLQIVRSLEWGSODELEAEVRQENRKKOKK  
 PDKVTEARRAERSVSWKRETIIVQHEVGRPENEEDMYRRTCTWCGHETVEKXK"  
 261..884  
 /note="XPA: Region: XPA protein"  
 /db\_xref="CDD:pfam01286"

## CDS

## misc\_feature

## ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1439;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
 |||||  
 DB 808 GGTCCATCTCATGTTGATG 789

RESULT 9  
 AF503166 24994 bp DNA linear PRI 25-APR-2002  
 LOCUS AF503166  
 DEFINITION Homo sapiens xeroderma pigmentosum, complementation group A (XPA)  
 gene, complete cds.

ACCESSION AF503166  
 VERSION AF503166.1 GI:20303097  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

## REFERENCE

1 (bases 1 to 24994)  
 Rieder, M.J., Branau, A.C., Montoya, M.A., Chung, M.-W., Nguyen, C.P.,  
 Nguyen, D.A., Livingston, R.J., Poel, C.L., Robertson, P.D.,  
 Schackwitz, W.S., Sherwood, J.K., Wittek, L.A. and Nickerson, D.A.

## TITLE

Direct Submission  
 Submitted (12-APR-2002) Genome Sciences, University of Washington,  
 1705 NE Pacific, Seattle, WA 98195, USA

## COMMENT

To cite this work please use: NIEHS-SNPs, Environmental Genome  
 Project, NIEHS S15478, Department of Genome Sciences, Seattle, WA  
 (URL: <http://esg.gs.washington.edu>).

## FEATURES

## source

1..24994  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 95..181  
 /rpt\_family="Alu"  
 repeat\_region

```
/rpt_type=dispersed
repeat_region 210..431
/rpt_family="Alu"
/rpt_type=dispersed
variation 259
/frequency="0.13"
/replacement="t"
variation 395
/frequency="0.15"
/replacement="g"
repeat_region 433..493
/rpt_family="L1"
/rpt_type=dispersed
variation 535
/frequency="0.01"
/replacement="a"
variation 600
/frequency="0.16"
/replacement="g"
variation 701
/frequency="0.01"
/replacement="c"
variation 706
/frequency="0.15"
/replacement="t"
repeat_region 1013..1076
/rpt_family="MIR"
/rpt_type=dispersed
variation 1168
/frequency="0.14"
/replacement="t"
repeat_region 1169..1381
/rpt_family="MIR"
/rpt_type=dispersed
variation 1428
/frequency="0.01"
/replacement="t"
repeat_region 1536..1853
/rpt_family="L1"
/rpt_type=dispersed
variation 1595
/frequency="0.01"
/replacement="c"
gene <1643..24097
/gene="XPA"
join(<1643..1840,5207..5317,9329..9434,11710..11875,
13931..14048,23420..24097)
/gene="XPA"
/product="xeroderma pigmentosum, complementation group A"
variation 1665
/gene="XPA"
/frequency="0.38"
/replacement="a"
CDS join(1669..1840,5207..5317,9329..9434,11710..11875,
13931..14048,23420..23568)
/gene="XPA"
/codon_start=1
/product="xeroderma pigmentosum, complementation group A"
/protein_id="XAM18969.1"
/db_xref="GI:20303098"
/translation="MAADGALPEAALDEQPAELPASVASIERKQRLMLRQRLA
APPSATIAAATGMAVKAAPKIDTGGFTLEEEBEEKIGKVHQPGVMEFDY
VIECEGKEFMSYLMNHFDLPDCNRDADDKHLITTEAKOEYLKCDLEREP
PIKEITVKRNPHSOMGDMKLYLKLOIVKRSLEWGSQALFEAKREVRQNRKQOK
FKKVKELRAVRAVSSVWKRETTIVHQHEFYEPENLEDDMRKTCMGHELTTERKM"
variation 1934
/gene="XPA"
/frequency="0.15"
/replacement="c"
variation 2044
/gene="XPA"
/frequency="0.13"
/replacement="g"

repeat_region 2182..2486
/rpt_family="CR1"
/rpt_type=dispersed
variation 2195
/gene="XPA"
/frequency="0.08"
/replacement="a"
variation 2424
/gene="XPA"
/frequency="0.02"
/replacement="g"
variation 2680
/gene="XPA"
/frequency="0.15"
/replacement="a"
variation 2903
/gene="XPA"
/frequency="0.16"
/replacement="g"
variation 3134
/gene="XPA"
/frequency="0.15"
/replacement="t"
variation 3173
/gene="XPA"
/frequency="0.02"
/replacement="g"
variation 3484
/gene="XPA"
/frequency="0.03"
/replacement="c"
variation 3509
/gene="XPA"
/frequency="0.05"
/replacement="g"
repeat_region 3643..3706
/rpt_family="MIR"
/rpt_type=dispersed
variation 3738..3742
/gene="XPA"
/frequency="0.85"
/replacement="t"
repeat_region 3785..4598
/rpt_family="L1"
/rpt_type=dispersed
variation 3818
/gene="XPA"
/frequency="0.11"
/replacement="a"
variation 4188
/gene="XPA"
/frequency="0.05"
/replacement="a"
variation 4487
/gene="XPA"
/frequency="0.04"
/replacement="t"
repeat_region 4607..4694
/rpt_family="MER2_type"
/rpt_type=dispersed
variation 4651
/gene="XPA"
/frequency="0.02"
/replacement="t"
variation 4710
/gene="XPA"
/frequency="0.02"
/replacement="t"
variation 4711..4714
/gene="XPA"
/frequency="0.06"
/replacement="t"
repeat_region 4738..5108
```

```

variation      /rpt_family="L1"
                /rpt_type="dispersed
4788            /gene="XPA"
                /frequency="0.02"
                /replace="c"
variation      4839
                /gene="XPA"
                /frequency="0.16"
                /replace="t"
variation      5272..5274
                /gene="XPA"
                /note="deletion results in removal of glutamate residue
                from amino acid sequence"
                /frequency="0.01"
                /replace="n"
variation      5522
                /gene="XPA"
                /frequency="0.01"
                /replace="a"
                /replace="a"
repeat_region  5574..5668
                /rpt_family="CR1"
                /rpt_type="dispersed
misc_feature   5815..5980
                /gene="XPA"
                /note="Region not scanned for variation"
repeat_region  5830..6001

Query Match    100.0%; Score 20; DB 9; Length 24994;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCACTCAGTGGATG 20
    |||||
Db 23489 GGTCCACTCAGTGGATG 23470

RESULT 10
AL445531      111345 bp   DNA   linear   PRI 01-MAR-2001
LOCUS         Human DNA sequence from clone RP11-54606 on chromosome 9, complete
DEFINITION    sequence.
ACCESSION     AL445531
VERSION       AL445531.10 GI:13234986
KEYWORDS      HTG.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 111345)
REFERENCE     Hammond S.
              Direct Submissiion
              Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
              requests: clonerequest@sanger.ac.uk
              On Mar 5, 2001 this sequence version replaced gi:11184378.
              During sequence assembly data is compared from overlapping clones.
              Where differences are found these are annotated as variations
              together with a note of the overlapping clone name. Note that the
              variation annotation may not be found in the sequence submission
              corresponding to the overlapping clone, as we submit sequences with
              only a small overlap as described above.
              This sequence was finished as follows unless otherwise noted: all
              regions were either double-stranded or sequenced with an alternate
              chemistry or covered by high quality data (i.e., phred quality >=
              30); an attempt was made to resolve all sequencing problems, such
              as compressions and repeats; all regions were covered by at least
              one plasmid subclone or more than one M13 subclone; and the
              assembly was confirmed by restriction digest. The following
              abbreviations are used to associate primary accession numbers given
              in the feature table with their source databases: Em: EMBL; Sw:
              SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
              database can be found at

```

```

FEATURES
  source
    http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
    was generated from part of bacterial clone contigs of human
    chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
    Group. Further information can be found at
    http://www.sanger.ac.uk/HGP/Chr9
    RP11-54606 is from the library RPCT-11.2 constructed by the group
    of Pieter de Jong. For further details see
    http://www.chori.org/bacpac/home.htm
    VECTOR: pBACe3.6
    IMPORTANT: This sequence is not the entire insert of clone
    RP11-54606. It may be shorter because we sequence overlapping
    sections only once, except for a 100 base overlap.
    The true left end of clone RP11-23B15 is at 111246 in this
    sequence. The true right end of clone RP11-24A49 is at 100 in this
    sequence.
    location/Qualifiers
    1..111345
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="9"
      /clone="RP11-54606"
      /clone_11b="RPCT-11.2"
      /note="MER5B repeat: matches 2..151 of consensus"
      2143..2375
        /note="R1gger2a repeat: matches 1..212 of consensus"
      2376..2671
        /note="Aluv repeat: matches 1..236 of consensus"
      2672..2836
        /note="R1gger2a repeat: matches 212..434 of consensus"
      3484..3661
        /note="MIR repeat: matches 28..250 of consensus"
      4208..4283
        /note="MIR repeat: matches 181..254 of consensus"
      6161..6299
        /note="MIR repeat: matches 79..233 of consensus"
      6933..7250
        /note="Aluv repeat: matches 1..310 of consensus"
      7601..7647
        /note="R1MCA repeat: matches 6807..6852 of consensus"
      7648..7954
        /note="Aluv repeat: matches 1..307 of consensus"
      7955..8098
        /note="R1MCA repeat: matches 6852..6999 of consensus"
      8266..8937
        /note="R1MCA repeat: matches 7297..7976 of consensus"
      8942..9568
        /note="R1MCA repeat: matches 6854..7538 of consensus"
      9621..9741
        /note="R1MCA repeat: matches 7625..7757 of consensus"
      9742..10054
        /note="Aluv repeat: matches 1..296 of consensus"
      10055..10134
        /note="R1MCA repeat: matches 7757..7833 of consensus"
      13726..13782
        /note="L2 repeat: matches 2654..2710 of consensus"
      13932..14026
        /note="MIR repeat: matches 37..141 of consensus"
      16174..16458
        /note="Aluv repeat: matches 1..306 of consensus"
      16989..17178
        /note="L2 repeat: matches 2223..2419 of consensus"
      17214..17390
        /note="L2 repeat: matches 1949..2126 of consensus"
      17428..17503
        /note="L2 repeat: matches 1797..1869 of consensus"
      17706..18041
        /note="W1TID repeat: matches 57..382 of consensus"
      19266..19365
        /note="50 copies 2 mer aa 62% conserved"
      19416..19742
        /note="W1TID repeat: matches 204..505 of consensus"

```

repeat_region	19743. .19906	/note="AluSx repeat: matches 139. .297 of consensus"
repeat_region	19907. .20214	/note="AluSg repeat: matches 1. .308 of consensus"
repeat_region	20215. .20351	/note="AluSx repeat: matches 1. .139 of consensus"
repeat_region	20352. .20548	/note="MLTID repeat: matches 1. .204 of consensus"
repeat_region	20757. .20833	/note="U2 repeat: matches 2576. .2661 of consensus"
repeat_region	20954. .21262	/note="AluSg repeat: matches 6. .303 of consensus"
repeat_region	21263. .21416	/note="AluSc repeat: matches 132. .285 of consensus"
repeat_region	21417. .21456	/note="20 copies 2 mer ta 92% conserved"
repeat_region	21540. .21839	/note="AluSp repeat: matches 1. .302 of consensus"
repeat_region	22282. .22443	/note="Charlie3 repeat: matches 2531. .2685 of consensus"
repeat_region	23387. .23579	/note="AluI4 repeat: matches 4077. .4280 of consensus"
repeat_region	23572. .23757	/note="U1M4 repeat: matches 4416. .4608 of consensus"
repeat_region	23801. .23887	/note="MER44 repeat: matches 637. .724 of consensus"
repeat_region	23897. .24110	/note="U1M53 repeat: matches 5239. .6095 of consensus"
misc_feature	26344. .27022	/note="Opc island"
repeat_region	27109. .27321	/evidence=not_experimental
repeat_region	27997. .28057	/note="MIR repeat: matches 9. .250 of consensus"
repeat_region	28059. .28280	/note="U1M55 repeat: matches 5098. .5164 of consensus"
repeat_region	28313. .28395	/note="AluI0 repeat: matches 1. .220 of consensus"
repeat_region	30327. .30625	/note="AluU repeat: matches 209. .291 of consensus"
repeat_region	30684. .31515	/note="U1M55 repeat: matches 5152. .5456 of consensus"
repeat_region	31516. .32221	/note="HERV1.48 repeat: matches 1103. .1927 of consensus"
repeat_region	32463. .32508	/note="U1M55 repeat: matches 5463. .6176 of consensus"
repeat_region	32518. .32784	/note="MER5A repeat: matches 61. .106 of consensus"
repeat_region	32877. .33055	/note="U1M61 repeat: matches 6056. .6325 of consensus"
repeat_region	34484. .34613	/note="MER5B repeat: matches 5. .174 of consensus"
repeat_region	34819. .35010	/note="U2 repeat: matches 2576. .2710 of consensus"
repeat_region	35519. .35707	/note="MIR repeat: matches 5. .193 of consensus"
repeat_region	35920. .36029	/note="U2 repeat: matches 2006. .2204 of consensus"
repeat_region	36428. .37325	/note="U2 repeat: matches 2627. .2750 of consensus"
repeat_region	37597. .37958	/note="U1M58 repeat: matches 5258. .6171 of consensus"
repeat_region	38790. .38968	/note="HER1B repeat: matches 1. .364 of consensus"
repeat_region	39203. .39497	/note="AluI0 repeat: matches 1. .175 of consensus"
repeat_region	41276. .41626	/note="AluSc repeat: matches 1. .296 of consensus"
repeat_region	41741. .42056	/note="U1M53 repeat: matches 5800. .6155 of consensus"
repeat_region	42238. .42458	/note="AluI0 repeat: matches 1. .310 of consensus"
repeat_region	42238. .42458	/note="MER30 repeat: matches 1. .230 of consensus"

Query Match	Best Local Similarity	100.0%;	Score 20;	DB 9;	Length 111345;
Matches	20;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;		
Query Match	Best Local Similarity	100.0%;	Score 20;	DB 9;	Length 111345;
Matches	20;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;		
Db	5047	GGTCATACATCATGTTGATG	5066		
Result 11	AL442130/c	187079 bp	DNA	linear	HTG 10-JUL-2001
LOCUS	AL442130	187079 bp	DNA	linear	HTG 10-JUL-2001
DEFINITION	Homo sapiens chromosome 9 clone RP11-24E19, 7 unordered pieces.				
ACCESSION	AL442130				
VERSION	AL442130.2	GI:10716422			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_CANCELLED.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	Plumb, B.			
AUTHORS	Direct Submission				
TITLE	Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,				
JOURNAL	CB10 ISA, UK. E-mail enquirer: humquery@sanger.ac.uk Clone				
COMMENT	requests: clonerequest@sanger.ac.uk				
	On Oct 7, 2000 this sequence version replaced gi:10278028.				
	----- Genome Center				
	Center: Sanger Centre				
	Web site: http://www.sanger.ac.uk				
	Contact: humquery@sanger.ac.uk				
	----- Project Information				
	Center project name: ba24E19				
	----- Summary Statistics				
	Assembly program: XGAP4; version 4.5				
	Sequencing vector: plasmid; 108752; 100% of reads				
	Chemistry: Dye-terminator Big Dye; 100% of reads				
	Consensus quality: 184692 bases at least Q40				
	Consensus quality: 185645 bases at least Q30				
	Consensus quality: 186064 bases at least Q20				
	Insert size: 186479; sum-of-contigs				
	Insert size: 117845; 70.8% error; agarose-fp				
	Insert coverage: 5.87x in Q20 bases; sum-of-contigs Quality				
	coverage: 10.17x in Q20 bases; agarose-fp				
	-----				
	* NOTE: This is a 'working draft' sequence. It currently				
	* consists of 7 contigs. The true order of the pieces				
	* is not known and their order in this sequence record is				
	* arbitrary. Gaps between the contigs are represented as				
	* runs of N, but the exact sizes of the gaps are unknown.				
	* This record will be updated with the finished sequence				

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence



```

* as soon as it is available and the accession number will
* be preserved.
1 43260: contig of 43260 bp in length
* 43261 43360: gap of 100 bp
* 43361 89644: contig of 46284 bp in length
* 89645 89744: gap of 100 bp
* 89745 94532: contig of 4788 bp in length
* 94533 94632: gap of 100 bp
* 94633 96636: contig of 2004 bp in length
* 96637 96736: gap of 100 bp
* 96737 129317: contig of 32581 bp in length
* 129318 129417: gap of 100 bp
* 129418 166289: contig of 36872 bp in length
* 166290 166390: gap of 100 bp
* 166390 187079: contig of 20690 bp in length.
FEATURES
    source
        1..187079
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="9"
            /clone="RP11-24B19"
            /clone_11b="RP11-11.1"
            1..43260
                /note="assembly fragment:01729
                fragment chain:1
                clone_end:17
                vector_side:left"
            43361..89644
                /note="assembly fragment:01187
                fragment chain:1"
            89745..94532
                /note="assembly fragment:01228
                fragment chain:1"
            94633..96636
                /note="assembly fragment:01200"
            96737..129317
                /note="assembly fragment:01705"
            129418..166289
                /note="assembly fragment:01977
                fragment chain:2"
            166390..187079
                /note="assembly fragment:01901
                fragment chain:2
                clone_end:886
                vector_side:right"
    misc_feature
        100.0%; Score 20; DB 2; Length 187079;
        Best Local Similarity 100.0%; Pred. No. 21;
        Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY
        1 GGTCCATCTACTGTTGATG 20
        |||||
        Db 78024 GGTCCATCTACTGTTGATG 78005

RESULT 12
XELXPACB2/c 1095 bp mRNA linear VRT 08-JAN-2003
LOCUS XELXPACB2 1095 bp mRNA linear VRT 08-JAN-2003
DEFINITION Xenopus laevis mRNA for xpacx2 protein, complete cds.
ACCESSION D31895
VERSION D31895.1 GI:505149
KEYWORDS
    ORGANISM
        Xenopus laevis (African clawed frog)
        Xenopus laevis
        Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
        Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
        Xenopodinae; Xenopus.
        1 (bases 1 to 1095)
REFERENCE
    AUTHORS
        Shimamoto,T., Kohno,K., Tanaka,K. and Okada,Y.
        Molecular cloning of human XPAC gene homologs from chicken, Xenopus
        laevis and Drosophila melanogaster

```

```

JOURNAL Biochem. Biophys. Res. Commun. 181 (3), 1231-1237 (1991)
MEDLINE 92109732
PUBMED 1764072
REFERENCE
    AUTHORS
        Shimamoto,T.
    TITLE
        Direct Submission
    JOURNAL
        Submitted (22-JUN-1994) Takuya Shimamoto, Osaka University Medical
        School, Department of Anatomy and Cell Biology, 2-2 Yamada-oka,
        Suita, Osaka 565, Japan (Tel:81-6-6879-5111(ex.3211),
        Fax:81-6-6879-3219)
FEATURES
    source
        1..1095
            /organism="Xenopus laevis"
            /mol_type="mRNA"
            /db_xref="taxon:8355"
            /clone="XPACX2"
            1..1095
                /gene="XPAC"
                1..250
                    /gene="XPAC"
                251..1048
                    /gene="XPAC"
            /note="Xeroderma Pigmentosus Group A complementing protein"
            /citation=[1]
            /codon_start=1
            /evidence=experimental
            /product="xpacx2 protein"
            /protein_id="BA06693.1"
            /db_xref="GI:505150"
            /translation="MEPEPEQANKEEKILISAARAKIERNRORALMLROARLAVRP
            YPTGEGISTVAKAPKVIDSGGPFIEEAEAEQHVENVVRQGPVLECDYLICEGCG
            DPMSTYLSNHDILAVCDSCRADEKRLITTEAKQETLDDCIDREPLKFTLKK
            NPHNTHWGDMLYLAQVTKRSLEWSEELBEAKERKRNDRMDKKFDKVKEL
            RRTVRSLSMKKEASGHQHEYPGEHVEEDSYKTCITCGYEMNYEKM"
            1049..1095
                /gene="XPAC"
            1071..1076
                /gene="XPAC"
            /gene="XPAC"
    polyA_signal
        3'UTR
        ORIGIN
            92.0%; Score 18.4; DB 5; Length 1095;
            Best Local Similarity 95.0%; Pred. No. 1.1e+02;
            Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
    QY
        1 GGTCCATCTACTGTTGATG 20
        |||||
        Db 969 GGTCCATCTACTGTTGATG 950

RESULT 13
XELXPACB 1277 bp mRNA linear VRT 08-JAN-2003
LOCUS XELXPACB 1277 bp mRNA linear VRT 08-JAN-2003
DEFINITION Xenopus laevis mRNA for xpacx1 protein, complete cds.
ACCESSION D31894
VERSION D31894.1 GI:505147
KEYWORDS
    ORGANISM
        Xenopus laevis (African clawed frog)
        Xenopus laevis
        Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
        Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
        Xenopodinae; Xenopus.
        1 (bases 1 to 1277)
REFERENCE
    AUTHORS
        Shimamoto,T., Kohno,K., Tanaka,K. and Okada,Y.
        Molecular cloning of human XPAC gene homologs from chicken, Xenopus
        laevis and Drosophila melanogaster
        Biochem. Biophys. Res. Commun. 181 (3), 1231-1237 (1991)
    MEDLINE
        92109732
    PUBMED
        1764072
    REFERENCE
        2 (bases 1 to 1277)
        Shimamoto,T.
        Direct Submission
        Submitted (22-JUN-1994) Takuya Shimamoto, Osaka University Medical
        School, Department of Anatomy and Cell Biology, 2-2 Yamada-oka,
        Suita, Osaka 565, Japan (Tel:81-6-6879-5111(ex.3211),
        Fax:81-6-6879-3219)

```

Suita, Osaka 565, Japan (Tel:81-6-6879-5111(ex.3211),  
Fax:81-6-6879-3219)

FEATURES  
source location/Qualifiers

1..1277  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="XPACX1"  
1..1277  
/gene="XPAC"  
1..67  
/gene="XPAC"  
68..871  
/gene="XPAC"  
/note="Xeroderma Pigmentosu Group A complementing protein"  
/citation=[1]  
/evidence=experimental  
/product="xpac1 protein"  
/protein\_id="BAA06692.1"  
/db\_xref="GI:505148"  
/translation="MEPEPEPEQANKEEKEKLSAARAKTERNRQRLMLRQRLAC  
RPPYPTGEGISTVAPKPVIDSGGFFIEEBAEQHVENVAPQPPVLECDYLIEEC  
GDFMDSYLSNHFDLAVDCSDAEKHLITRTAEQYELKDCDIDKREYLFIL  
KKNPHNTGMDKLYLKAQVIRSLVMSGEALBEAKVRKNDKMKOKKFDKKVK  
ELRRVTRSSLMKEASGHQHEYGPEHVEDESYKKTCTICGYENMYEKM"  
872..1277  
/gene="XPAC"  
894..899  
/gene="XPAC"  
1257..1262  
/gene="XPAC"  
1277  
/gene="XPAC"

3'UTR  
polyA\_signal  
polyA\_signal  
polyA\_site  
/gene="XPAC"

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 1277;  
Best Local Similarity 95.0%; Pred. No. 1.1e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTATG 20  
|||||  
DB 792 GGTCCATCTCATGTGTATG 773

RESULT 14  
AL732555 157653 bp DNA linear ROD 25-JUL-2002  
LOCUS AL732555/c  
DEFINITION Mouse DNA sequence from clone RP23-124119 on chromosome 4, complete  
sequence.  
ACCESSION AL732555  
VERSION AL732555.7 GI:21998253  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 157653)  
Lovel, J.  
Direct Submission  
Submitted (25-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Jul 26, 2002 this sequence version replaced gi:21955637.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----

COMMENT  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPep; information on the WormPep  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-124119 is  
from the RPI-23 Mouse PAC Library  
constructed by the group of Peter de Jong.  
For further details see http://www.choxi.org/bacpac/home.htm  
VECTOR: pBAC3.6.

FEATURES  
source location/Qualifiers

1..157653  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone="RP23-124119"  
/clone\_11b="RPI-23"

ORIGIN

Query Match 92.0%; Score 18.4; DB 10; Length 157653;  
Best Local Similarity 95.0%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTATG 20  
|||||  
DB 143430 GGTCCATCTCATGTGTATG 143411

RESULT 15  
AC109694 199263 bp DNA linear HTG 22-SEP-2002  
LOCUS AC109694  
DEFINITION Rattus norvegicus clone CH230-200G12, \*\*\* SEQUENCING IN PROGRESS  
AC109694  
VERSION AC109694.5 GI:23195474  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 199263)  
AUTHORS Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Banstead, M., Benahmed, F.,  
Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Cavazos, I., Ceeasar, H., Centes, A.,  
Cardenas, V., Carter, K., Cavazos, J., Ceasar, H., Centes, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Chevaland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Dayla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,  
Gunaratne, P., Haaland, N., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Huljk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,



Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, M., Guenara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harrey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Holling, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorusshewa, L., Louised, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nat, L., Nankervill, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleme, O., Okunolu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Rutz, S.U., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabar, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Vales, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Wooley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, D., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Wernstock, G. and Gibbs, R.A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished  
2 (bases 1 to 227524)  
Worley, K.C.  
Direct Submission  
Submitted (04-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 227524)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 13, 2003 this sequence version replaced gi:24942467.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GKNS  
Center clone name: CH230-208J22  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 194004 bases at least Q40

Consensus quality: 197089 bases at least Q30  
Consensus quality: 199168 bases at least Q20  
Estimated insert size: 205215; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 227524 contig of 227524 bp in length.  
Location/Qualifiers  
1..227524  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-208J22"  
1..2135  
/note="wgs\_end\_extension  
clone\_end:"77"  
4585..5409  
/note="clone\_boundary  
clone\_end:"77"  
site:ECORI  
end\_sequence:RMBKP59TJB"  
226063..226880  
/note="clone boundary  
clone\_end:Sp6  
site:ECORI  
end\_sequence:RMBKP59TJB"

## ORIGIN

Query Match 92.0%; Score 18.4; DB 2; Length 227524;  
Best Local Similarity 95.0%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20  
|||||  
Db 58568 GGTCCTACTCATGTTGATG 58587

RESULT 17  
AP003607/c  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,  
PAC clone: P0004A09.  
ACCESSION  
AP003607.3 GI:15290074  
VERSION  
AP003607.3  
KEYWORDS  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1  
REFERENCE  
AUTHORS  
Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,  
Katayose, Y., Wu, J., Nilmura, Y., Cheng, Z., Nagamura, Y.,  
Antonio, B.A., Kanamori, H., Hoshikawa, S., Maekawa, M., Y.,  
Chien, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,  
Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,  
Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, T.,  
Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karsawa, W., Katagiri, S.,  
Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,  
Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M.,  
Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,  
Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,  
Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K.,

TITLE	The genome sequence and structure of rice chromosome 1
JOURNAL	Nature 420 (6913), 312-316 (2002)
	Yamagata, H., Yamane, H., Yoshiki, S., Yoshinaka, R., Yukuwa, K., Zhong, H., Iwana, H., Endo, T., Ito, H., Hahn, J. H., Kim, H. I., Eun, M. Y., Yano, M., Jiang, J. and Gojobori, T.

CDS  
39061..40536  
/gene="P0004A09.9"  
/note="contains ESTs C26603 (C12674), AU075575 (C12674)"  
/codon\_start=1  
/product="putative galicylate-induced glucosyltransferase"  
/protein\_id="BAB63773.1"  
/db\_xref="GI:15290080"  
/translation="NAIKDEOQPLHLPFPFLAPGLIPIADMALFARGVCTILT  
TPNAAVIRSAVDRANDSPRRNGGLAIETVPPDVLPGPFGSGTALTQDDRD  
PLIGRIIHEPDRYLSHHNDAAYVDSFRAADAAAEHCPGPGICTSVFARCT  
NSMLKNPLETAPDDPAVPLPLGLPHCEVLRSSQMDPKAPDMEKRFQSIDAQR  
SFEVFNSEFHELEPDYVEHYRTTIGRRVWLVGVPAIANKDVAVRGTSLSLSPADGYLR  
WLDAPRGSVVVVSFGTILSSFSFPAEMREIARGLDLSGNFVWVINGADADASEWMPBG  
FAELIAPRGESRLTIRGAPOMLILNHFPVGGFVTCGNSTLEAVTAGVPMVTPRY  
ADQFNEKLTILEVLEGVGVSMDPASKIENRRTVIGGVVGAIGRWVGDEEBEAI  
RKKATELVKAGALEKGGSSVDVGLIMDELMARCSNVN"  
complement(join(43936..43947,44152..44568))  
/gene="P0004A09.10"  
complement(join(43936..43947,44152..44568))  
/gene="P0004A09.10"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="BAB63774.1"  
/db\_xref="GI:15290081"  
/translation="MNTDARRARRPRTRMWGSAPAVEAEAAAMIRSGGGGGGGRG  
RLDDGDMDEGVDDNSTTARQWDLLEGVDDSDTAAARRRRLSSPPMPSVVLA  
SETATDIFPPPSHSSAHPLASPLPHRRLDLSVARPIIA"  
complement(join(45670..46125,46162..46395))  
/gene="P0004A09.11"  
complement(join(45670..46125,46162..46395))  
CDS  
Query Match  
Best Local Similarity 94.7%; Pred. No. 5e+02; Length 158826;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GGTCCATCTCATGTTGAT 19  
|||||  
Db 25634 GGTCCATCTCATGTTGAT 25616  
RESULT 18  
AP004367/c 197674 bp DNA linear HTG 27-NOV-2003  
LOCUS Oryza sativa (japonica cultivar-group) chromosome 1 clone P0696E01,  
DEFINITION \*\*\* SEQUENCING IN PROGRESS \*\*\*  
ACCESSION AP004367 BA000010  
VERSION AP004367.2 GI:21952912  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhacridae; Oryzaceae; Oryza.  
REFERENCE  
AUTHORS  
1 Sasaki,T., Matsumoto,T., Yamamoto,K., Sakata,K., Baba,T.,  
Katayose,Y., Wu,J., Nilmura,Y., Cheng,Z., Nagamura,Y.,  
Antonio,B.A., Kanamori,H., Hosokawa,S., Masukawa,K., Arikawa,K.,  
Chiden,Y., Hayashi,M., Okamoto,M., Ando,T., Aoki,H., Arika,K.,  
Hamada,M., Harada,C., Hijishita,S., Honda,M., Ichikawa,Y.,  
Idonuma,A., Iijima,M., Ikeda,M., Ikeda,M., Itoh,S., Itoh,T.,  
Kikuta,A., Kobayashi,N., Kono,I., Machita,K., Mebara,T.,  
Mizuno,H., Miyabayashi,T., Mukai,Y., Nagasaki,H., Nakashima,M.,  
Nakama,Y., Nakamichi,Y., Nakamura,M., Naito,N., Negishi,M.,  
Ohta,I., Ono,N., Saji,S., Sakai,K., Shibata,M., Shimokawa,T.,  
Shomura,A., Song,J., Takazaki,Y., Terasawa,K., Tsuji,K., Waki,K.,  
Yamagata,H., Yamane,H., Yoshiki,S., Yoshikawa,R., Yukawa,K.,  
Zhong,H., Iwama,H., Endo,T., Ito,H., Hahn,J.H., Kim,H.I., Eun,M.Y.,  
Yano,M., Jiang,J. and Gojobori,T.  
TITLE The genome sequence and structure of rice chromosome 1  
JOURNAL Nature 420 (6913), 312-316 (2002)  
MEDLINE 22337376  
PUBMED 12447438

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
2 (bases 1 to 197674)  
Sasaki,T., Matsumoto,T. and Yamamoto,K.  
Direct Submission  
Submitted (19-NOV-2001) Takuji Sasaki, National Institute of  
Agricultural Sciences, Rice Genome Research Program, Kamondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rsgp.dna.affrc.go.jp/  
Tel:81-298-38-7441, Fax:81-298-38-7468)  
On Jul 24, 2002 this sequence version replaced gi:17046147.  
NOTE: It currently consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces is believed  
to be correct as given, however the sizes of the gaps between them  
are based on estimates that have provided by the submitter. This  
sequence will be replaced by the finished sequence as soon as it is  
available and the accession number will be preserved.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 197674: contig of 197674 bp in length.  
Location/Qualifiers  
1. 197674  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="1"  
/clone="P0696E01"  
FEATURES  
source  
ORIGIN  
Query Match  
Best Local Similarity 94.7%; Pred. No. 5.1e+02; Length 197674;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GGTCCATCTCATGTTGAT 19  
|||||  
Db 41502 GGTCCATCTCATGTTGAT 41484  
RESULT 19  
AC115118/c 198448 bp DNA linear ROD 05-NOV-2003  
LOCUS Mus musculus BAC clone RP23-84N24 from 5, complete sequence.  
DEFINITION AC115118  
ACCESSION AC115118.3 GI:22138690  
VERSION AC115118.3 GI:22138690  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
1 (bases 1 to 198448)  
Abbot,S., Haakenson,W. and Doeber,A.  
The sequence of Mus musculus BAC clone RP23-84N24  
REFERENCE  
AUTHORS  
2 (bases 1 to 198448)  
Wilson,R.  
Sequencing of Mus musculus  
Unpublished (2001)  
3 (bases 1 to 198448)  
McPherson,J.D. and Waterston,R.H.  
Direct Submission  
Submitted (14-MAR-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 198448)  
McPherson,J.D. and Waterston,R.H.  
Direct Submission  
Submitted (28-MAY-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 198448)  
AUTHORS McPherson,J.D. and Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (08-AUG-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
REFERENCE 6 (bases 1 to 198448)  
AUTHORS Wilson,R.  
TITLE Direct Submission  
JOURNAL Submitted (05-NOV-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Aug 8, 2002 this sequence version replaced gi:21218509.  
-----  
Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: MUGSC  
Web site: http://genome.wustl.edu  
Contact: submissions@watson.wustl.edu  
-----  
Summary Statistics  
Center project name: M\_BA0084N24  
-----  
NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.  
  
This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.  
  
MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
http://genome.wustl.edu  
  
SOURCE INFORMATION:  
The RPT-23 BAC library has been constructed by Kazutoyo Osegawa  
and Minako Tateno in the laboratory of Pieter de Jong  
(http://www.chori.org) from female C57BL/6J mouse kidney and/or  
brain genomic DNA. The clone and detailed information can be  
obtained from Research Genetics, Inc. (http://www.resgen.com) or  
Pieter de Jong and coworkers at http://www.chori.org  
  
NEIGHBORING SEQUENCE INFORMATION:  
This sequence is the entire insert of the clone. This clone is  
overlapped by AC122492:  
Location/Qualifiers  
1.198448  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="5"  
/map="5"  
/clone="RP23-84N24"  
/clone\_1lb="RPT-23"  
repeat\_region  
673..857  
/rpt\_family="B2"  
repeat\_region  
1952..2649  
/rpt\_family="L1"  
repeat\_region  
4688..5150  
/rpt\_family="L1"  
repeat\_region  
5186..5680  
/rpt\_family="L1"  
repeat\_region  
6562..7050  
/rpt\_family="L1"  
repeat\_region  
9091..9435  
/rpt\_family="MALR"  
repeat\_region  
10463..10982  
/rpt\_family="ERVK"

repeat\_region 11751..12096  
/rpt\_family="MALR"  
repeat\_region 1411..14209  
/rpt\_family="L1"  
repeat\_region 14210..14435  
/rpt\_family="L1"  
repeat\_region 14436..14620  
/rpt\_family="B2"  
tRNA complement(14542..14615)  
/product="tRNA-Ser"  
/note="Likely pseudogene (MM Sc=42.36 / Sec struct  
Sc=-22.25)"  
repeat\_region 14621..14951  
/rpt\_family="L1"  
repeat\_region 15355..16585  
/rpt\_family="L1"  
repeat\_region 16584..16622  
/rpt\_family="L1"  
repeat\_region 16623..17018  
/rpt\_family="MALR"  
repeat\_region 17019..18064  
/rpt\_family="L1"  
repeat\_region 18061..18488  
/rpt\_family="L1"  
repeat\_region 18477..18737  
/rpt\_family="L1"  
repeat\_region 18737..19336  
/rpt\_family="L1"  
repeat\_region 19721..20369  
/rpt\_family="L1"  
repeat\_region 21192..21376  
/rpt\_family="L1"  
repeat\_region 21377..21803  
/rpt\_family="ERVK"  
repeat\_region 21804..22035  
/rpt\_family="L1"  
repeat\_region 22058..22317  
/rpt\_family="L1"  
repeat\_region 22354..23361  
/rpt\_family="L1"  
repeat\_region 23798..24434  
/rpt\_family="L1"  
repeat\_region 26562..26741  
/rpt\_family="B2"  
repeat\_region 26886..27097  
/rpt\_family="B2"  
repeat\_region 27372..27501  
/rpt\_family="B2"  
repeat\_region 27681..27791  
/rpt\_family="A1u"  
repeat\_region 27806..27857  
/rpt\_family="MER1B"  
repeat\_region 29422..29579  
/rpt\_family="B4"  
repeat\_region 30476..30932  
/rpt\_family="L1"  
repeat\_region 30931..31525  
/rpt\_family="L1"  
repeat\_region 31526..31857  
/rpt\_family="MALR"  
repeat\_region 31901..32233  
/rpt\_family="ERVK"  
repeat\_region 35194..36055  
/rpt\_family="ERVK"  
repeat\_region 36058..36391  
/rpt\_family="ERVK"  
repeat\_region 36405..36491  
/rpt\_family="B4"  
repeat\_region 37726..38070  
/rpt\_family="MALR"  
repeat\_region 38166..38305  
/rpt\_family="MALR"  
repeat\_region 38306..38730



/rpt\_family="ERVX"  
repeat\_region 38731..38846  
/rpt\_family="MALR"  
repeat\_region 33695..33942  
/rpt\_family="Alu"  
repeat\_region 42175..42378  
/rpt\_family="MER1\_type"  
repeat\_region 42583..42754  
/rpt\_family="B2"  
repeat\_region 42758..42962  
/rpt\_family="B2"  
repeat\_region 43332..43672  
/rpt\_family="L1"  
repeat\_region 44595..44811  
/rpt\_family="B2"  
repeat\_region 45045..45215  
/rpt\_family="B2"  
repeat\_region 45267..45414  
/rpt\_family="Alu"  
repeat\_region 45570..45777  
/rpt\_family="MALR"  
repeat\_region 47156..47217  
/rpt\_family="B2"  
repeat\_region 47218..47556  
/rpt\_family="MALR"  
repeat\_region 47557..47670  
/rpt\_family="B2"

Query Match 87.0%; Score 17.4; DB 10; Length 198448;  
Best Local Similarity 94.7%; Pred. No. 5.1e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCGATGCTGATGCTGATG 20  
Db 102771 GTCGATGCTGATGCTGATG 102753

RESULT 20  
LOCUS HUAC004682 189134 bp DNA linear PRI 30-OCT-2002  
DEFINITION Homo sapiens Chromosome 16 BAC clone CIT987SK-A-259H10, complete  
sequence.  
ACCESSION AC004682  
VERSION AC004682.1 GI:3337388  
KEYWORDS HTG.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS  
Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,  
Fuhrmann,J., Mason,T., Crosby,M.L., Barnstead,M., Cronin,U.,  
Deslattes Maye,A., Cao,Y., Xu,R.X., Kang,H.T., Mitchell,S.,  
Eichler,E.E., Harris,P.C., Venter,J.C. and Adams,M.D.  
Mammalian; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 189134)  
Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,  
Fuhrmann,J., Mason,T., Crosby,M.L., Barnstead,M., Cronin,U.,  
Deslattes Maye,A., Cao,Y., Xu,R.X., Kang,H.T., Mitchell,S.,  
Eichler,E.E., Harris,P.C., Venter,J.C. and Adams,M.D.  
Genome duplications and other features in 12 Mb of DNA sequence  
from human chromosome 16p and 16q  
Genomics 60 (3), 295-308 (1999)  
JOURNAL MEDLINE  
PUBMED 99425270  
10493829  
TITLE (bases 1 to 189134)  
AUTHORS Adams,M.D., Loftus,B.J., Zhou,L., Crosby,M., Fuhrmann,J.,  
Mason,T.M., Brandon,R., Kim,U.J., Kerlavage,A.R. and Venter,J.C.  
Homo sapiens chromosome 16 BAC clone CIT987SK-A-259H10  
Unpublished  
3 (bases 1 to 189134)  
AUTHORS Adams,M.D. and Loftus,B.J.  
TITLE Direct Submission  
JOURNAL Submitted (12-MAY-1998) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, Email:  
bjloftus@tigr.org  
4 (bases 1 to 189134)  
REFERENCE  
AUTHORS Adams,M.D.  
TITLE Direct Submission

JOURNAL  
COMMENT  
Submitted (24-JUL-1998) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
On Jul 24, 1998 this sequence version replaced gi:3249114.  
Address all correspondence to: Mark Adams The Institute for Genomic  
Research 9712 Medical Center Dr, Rockville, MD 20850, USA e-mail  
address: humgen@tigr.org. The orientation of the sequence is from  
SP6 end to 3' end. Genes were identified by a combination of five  
methods including: XGRAIL (available by anonymous ftp from  
archur.epm.ornl.gov), GeneFinder (Phil Green, University of  
Washington), GenScan (Chris Burge,  
http://genomic.stanford.edu/~chris/GENSCANW.html) searches of the  
complete sequence against a peptide database, and the Human gene  
Index database at TIGR (http://www.tigr.org/tdb/hgi/hgi.html).  
Genes without peptide homology having spliced EST hits are termed  
'Unknown gene product'. Genes encoding tRNAs are predicted by  
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).

FEATURES  
SOURCE  
Location/Qualifiers  
1..189134  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/map="16q22.2"  
/clone="A-259H10"  
complement(133825..>150585)  
/gene="A-259H10.1"  
complement(join(133825..134330,135666..135787,  
137233..137328,137818..137943,138136..138270,  
138361..138516,138944..139098,139195..139403,  
140626..140738,141119..141228,141335..141511,  
141606..141716,142117..142257,142612..142797,  
142955..143141,143532..143669,145131..145243,  
145550..145657,146146..146307,146856..147011,  
147463..147539,147698..147816,147951..148098,  
148962..149066,149734..149921,150263..>150585))  
/gene="A-259H10.1"  
133855..133981  
EST165365, Chr. 16, Homo sapiens"  
/note="39391, EST165365, Chr. 16, Homo sapiens"  
/db\_xref="dbSTS:G25538"  
133855..133981  
/note="40992, SHGC-32168, Chr. 16, Homo sapiens"  
/db\_xref="dbSTS:G27137"  
complement(join(134246..134330,135666..135787,  
137233..137328,137818..137943,138136..138270,  
138361..138516,138944..139098,139195..139403,  
140626..140738,141119..141228,141335..141511,  
141606..141716,142117..142257,142612..142797,  
142955..143141,143532..143669,145131..145243,  
145550..145657,146146..146307,146856..147011,  
147463..147539,147698..147816,147951..148098,  
148962..149066,149734..149921,150263..150585))  
/gene="A-259H10.1"  
/codon\_start=1  
/product="pre-mRNA splicing factor (PRP16) (KIAA0224)"  
/protein\_id="AAC27431.1"  
/db\_xref="GI:3337389"

gene  
mRNA  
complement(join(133825..134330,135666..135787,  
137233..137328,137818..137943,138136..138270,  
138361..138516,138944..139098,139195..139403,  
140626..140738,141119..141228,141335..141511,  
141606..141716,142117..142257,142612..142797,  
142955..143141,143532..143669,145131..145243,  
145550..145657,146146..146307,146856..147011,  
147463..147539,147698..147816,147951..148098,  
148962..149066,149734..149921,150263..>150585))  
/gene="A-259H10.1"  
133855..133981  
EST165365, Chr. 16, Homo sapiens"  
/note="39391, EST165365, Chr. 16, Homo sapiens"  
/db\_xref="dbSTS:G25538"  
133855..133981  
/note="40992, SHGC-32168, Chr. 16, Homo sapiens"  
/db\_xref="dbSTS:G27137"

STS  
CDS  
complement(join(134246..134330,135666..135787,  
137233..137328,137818..137943,138136..138270,  
138361..138516,138944..139098,139195..139403,  
140626..140738,141119..141228,141335..141511,  
141606..141716,142117..142257,142612..142797,  
142955..143141,143532..143669,145131..145243,  
145550..145657,146146..146307,146856..147011,  
147463..147539,147698..147816,147951..148098,  
148962..149066,149734..149921,150263..150585))  
/gene="A-259H10.1"  
/codon\_start=1  
/product="pre-mRNA splicing factor (PRP16) (KIAA0224)"  
/protein\_id="AAC27431.1"  
/db\_xref="GI:3337389"

translation="MGDTSEDAASIHRLGTDLDCCVGLICKSKSAASEOHVFPAPAP  
RPSLLGLDILASIKRREEREDGDEKKSVSXKMEESKDDOKOAAEEGSGOAG  
NIRKORHYRSARVETPSHPGVSSEFPERSRQRERREHGVYASSGEKKKKEKR  
DRDYDRKDRDRDRSRSSRSERDGSERSRSTRDRSDVRKGYDDVPLPTPSKYNE  
DSGYSSRSQWESPSPSPSYRDSRGRSEEGSEFTEERQWEDQCAQADWYMEQY  
WADDRHILGSPPLSRGRSEEGSEFTEERQWEDQCAQADWYMEQY  
EFHNLVASSSDVYRRRQHHKOKRISQROINDENRQMETNRLTSGVHRL  
VDDEFEEDNAKVKHMTNINIVPPLDGIIVTQBPVIVKDATSDLAITARKGSOT  
VKRHEQKERKKAQHKEWELGTLGDMVKKEEPEPKAVTEGKQDYREQPADH  
MKRSEASSERPAKKSILRQRYLPIFAVQDELTLTRDNSIVIVGEGTSQKTLT  
OYLHEDGYTDMGICGTPRRVAAVSAAKRSSEMGNLGEVATLREEDTSENTL  
IKYHEDGYLRESLEADLDHYSAIMDEAHERSLNDVLEGLREVARSDDKLTIV  
TSATMDAKFAFAGNVPIPHIPIRTPPVDVLEFSTQPEDVEEAVVQSLGVHSGAP  
GDILTFMGQEDIEVTSQIVYEHLEEMALAVLPIVQSLPBDLAKIFOKAPDGV  
RKCIVATNIATSLTVDGIMFVDSGCKLKVFPRIQMDLQIPIPSQANANQBSGR  
AGRTPGQCFRLYTOSAYKNELTTTVEIQRNLAVNLILKLGVDLLOFHHMDP  
PPEDMLNMTYQWLIGALDVTGGLTSGRLMVEFPDLPALSKMLIVSCDMGCSSEIL



LIIVSLVPAIFPRPKGREESDOIKEKFAVESDHLTYLVNLYOMKNNNTIWCND  
HFIHAKAMKREVEBAQLKDIWVOORMSLASGTPMDIVKCIACAIFQAKLKIG  
EYVNIIRTGMPCHLPTSSIFGMYTPDYLIVHELVITKIEWOCYAVDSEVLATIG  
MEYSVKQAKSKQENRRRAKERAASAMEEMALAEOLRARROEQEKSPUSVRSKTI  
YTPGKEQCEEPTRTPARFGL"  
complement(<169670..>172858)  
/gene="A-259H10.2"  
complement(join(169670..170448,171532..171606,  
172366..172467,172774..>172858))  
/gene="A-259H10.2"  
complement(join(169670..170448,171532..171606,  
172366..172467,172774..>172858))  
/gene="A-259H10.2"  
/codon\_start=1  
/product="haploglobin-related protein precursor"  
/protein\_id="AAC27433.1"  
/db\_xref="GI:3337391"  
/translation="DLGAVYSLILNGRQLFALYSQNDVTDISDRFPKPEIANGYVE  
HLFRYCKNYRLRTGSDGYTLNDKQWINKAVGDKLPECAVCGKPKNPANPYQRI  
LGGHLDKAGSPFMOAKMVSHNLTGATLINQOMLTITAKNLFILHSEKATKDIAPT  
LTLVYKQVLVEIKYVLAHPNHYVDIGIKKOKYLVNERMPICLBKNAVYGRV  
GYSGWGSDNFKLTDHLKYVNLPVADQYDCTITHEGSTCPKPKAPKSPVQPIINE  
HTFCVMSKYQEDTCYGDAGSAFAVHDLDEDTWYAAILSPDKSCAVAEYGVYKVS  
IQHWYQKTIKEN"  
175654..175784  
/note="35342..WT-17126..Chr. 16, Homo sapiens"  
/db\_xref="dbSTS:G21504"  
complement(<185861..188871)  
/gene="A-259H10.3"  
complement(join(<185861..186639,187563..187637,  
188396..188497,188784..188871))  
/gene="A-259H10.3"  
complement(join(185861..186639,187563..187637,  
188396..188497,188784..>188865))  
/gene="A-259H10.3"  
/codon\_start=1  
/product="haploglobin"  
/protein\_id="AAC27432.1"  
/db\_xref="GI:3337390"  
/translation="ALGAVIALILNGQLFVAVDSNDVTDIADGCPKPEIANGYVEH  
SVRYCKNYKRLRTGSDGYTLNNEKQWINKAVGDKLPECAVCGKPKNPANPYQRI  
LGGHLDKAGSPFMOAKMVSHNLTGATLINQOMLTITAKNLFILHSEKATKDIAPT  
LTLVYKQVLVEIKYVLAHPNHYVDIGIKKOKYLVNERMPICLBKNAVYGRV  
GYSGWGSDNFKLTDHLKYVNLPVADQYDCTITHEGSTCPKPKAPKSPVQPIINE  
HTFCVMSKYQEDTCYGDAGSAFAVHDLDEDTWYAAILSPDKSCAVAEYGVYKVS  
IQHWYQKTIKEN"

ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 189134;  
Best Local Similarity 100.0%; Pred. No. 8.3e+02;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CCATCTCATGTTGATG 20  
|||  
DB 104856 CCATCTCATGTTGATG 104872

RESULT 21  
AC009160/c  
LOCUS AC009160.5 GI:8575975  
DEFINITION Homo sapiens chromosome 16 clone RP11-70B3, WORKING DRAFT SEQUENCE,  
27 unordered pieces.  
ACCESSION AC009160  
VERSION AC009160.5 GI:8575975  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 190856)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

2 (bases 1 to 190856)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jun 21, 2000 this sequence version replaced gi:7689973.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 433344  
Center clone name: RPCI-11\_70B3  
-----  
Summary Statistics  
Consensus quality: 162008 bases at least Q40  
Consensus quality: 174412 bases at least Q30  
Consensus quality: 177602 bases at least Q20  
Estimated insert size: 162600; agarose-fp estimation  
Estimated insert size: 188256; sum-of-contigs estimation  
Quality coverage: 5.83 in Q20 bases; agarose-fp estimation  
Quality coverage: 5.04 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 27 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1014: contig of 1014 bp in length  
\* 1015 1114: gap of unknown length  
\* 1115 2667: contig of 1553 bp in length  
\* 2668 4750: contig of 1983 bp in length  
\* 4751 4851: gap of unknown length  
\* 4851 6226: contig of 1376 bp in length  
\* 6227 6326: gap of unknown length  
\* 6327 7638: contig of 1312 bp in length  
\* 7639 7739: gap of unknown length  
\* 7739 8805: contig of 1067 bp in length  
\* 8806 8905: gap of unknown length  
\* 8906 10073: contig of 1168 bp in length  
\* 10074 10173: gap of unknown length  
\* 10174 11248: contig of 1075 bp in length  
\* 11249 11349: gap of unknown length  
\* 11349 12368: contig of 1020 bp in length  
\* 12369 12468: gap of unknown length  
\* 12469 13733: contig of 1265 bp in length  
\* 13734 13833: gap of unknown length  
\* 13834 15022: contig of 1189 bp in length  
\* 15023 15123: gap of unknown length  
\* 15123 16567: contig of 1445 bp in length  
\* 16568 16667: gap of unknown length  
\* 16668 18356: contig of 1689 bp in length  
\* 18357 18456: gap of unknown length  
\* 18457 22179: contig of 3723 bp in length  
\* 22180 22279: gap of unknown length  
\* 22280 25036: contig of 2757 bp in length  
\* 25037 25136: gap of unknown length  
\* 25137 31395: contig of 6259 bp in length  
\* 31396 31495: gap of unknown length  
\* 31496 39492: contig of 7997 bp in length  
\* 39493 39593: gap of unknown length  
\* 39593 48267: contig of 8674 bp in length  
\* 48267 48366: gap of unknown length  
\* 48367 60754: contig of 12388 bp in length  
\* 60755 60854: gap of unknown length  
\* 60855 71400: contig of 10546 bp in length  
\* 71401 71500: gap of unknown length  
\* 71501 81265: contig of 9765 bp in length  
\* 81266 81365: gap of unknown length

```

*      81366      95666: contig of 14301 bp in length
*      95667      95766: gap of unknown length
*      95767      108787: contig of 13021 bp in length
*      108788      108887: gap of unknown length
*      108888      127457: contig of 18569 bp in length
*      127457      127556: gap of unknown length
*      127556      147174: contig of 19618 bp in length
*      147175      147274: gap of unknown length
*      147275      168580: contig of 21306 bp in length
*      168581      168680: gap of unknown length
*      168681      190856: contig of 22176 bp in length.
FEATURES
    source
        1..190856
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="16"
        /clone="RP11-70E3"
        /clone_1ib="RP11 human BAC library 11"
ORIGIN
Query Match      85.0%; Score 17; DB 2; Length 190856;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      4 CCATCACTCATGTTGATG 20
Db      157563 CCATCACTCATGTTGATG 157547
RESULT 22
AC134076      253297 bp      DNA      linear      HTG 13-NOV-2002
LOCUS      Rattus norvegicus clone CH230-2F4, WORKING DRAFT SEQUENCE.
ACCESSION      AC134076.2 GI:24941326
VERSION      HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS      Rattus norvegicus (Norway rat)
SOURCE      Rattus norvegicus
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
      Rattus.
REFERENCE      1 (bases 1 to 253297)
AUTHORS      Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
      Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
      Barbarta,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
      Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
      Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
      Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
      Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
      Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
      Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.H.,
      Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
      Doumwaite,K.U., Draper,H., Dugan-Rocha,S., Durkin,K.U.,
      Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
      Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
      Gabrieli,J., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
      Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
      Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
      Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
      Homel,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.B.,
      Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
      Karlsson,E., Kelly,S., Khan,U., King,L., Kovvah,J., Kovar,C.,
      Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
      Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,W., Louised,H.,
      Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
      Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
      Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
      Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
      Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
      Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okunolu,G.,
      Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
      Peters,L., Pickens,R., Plimus,E., Pu,L.L., Quiles,M., Ren,Y.,

```

```

Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoochert,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,J., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 253297)
Worley,K.C.
Direct Submission
Submitted (22-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253297)
Worley,K.C.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23267443.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: TURX
Center clone name: CH230-2F4
----- Summary Statistics
Assembly program: phrap; version 0.990329
Consensus quality: 234461 bases at least Q40
Consensus quality: 236497 bases at least Q30
Consensus quality: 237862 bases at least Q20
Estimated insert size: 242145; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 253297: contig of 253297 bp in length.
FEATURES
    source
        1..253297
        /organism="Rattus norvegicus"
        /mol_type="genomic DNA"
        /db_xref="taxon:10116"
        /clone="CH230-2F4"
        /clone="CH230-2F4"
        /note="wgs_contig"
misc_feature

```

```

misc_feature      244939..246775
                    /note="wgs_contig"
misc_feature      251906..253297
                    /note="wgs_contig"

ORIGIN

Query Match      85.0%; Score 17; DB 2; Length 253297;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGTCCATCTCATGTTG 17
Db      113342 GGTCCATCTCATGTTG 113358

RESULT 23
AC094314/c
LOCUS
DEFINITION
2 unordered pieces.
AC094314      263730 bp      DNA      linear      HTG 09-MAY-2003
Rattus norvegicus clone CH230-3G10, *** SEQUENCING IN PROGRESS ***,
AC094314
AC094314.8 GI:30467383
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 263730)
Muzny,D,Marie,, Metzker,M,Dee,, Abramson,S,, Adams,C,, Alder,J,,
Allen,C,, Allen,H,, Alshrooks,S,, Amin,A,, Angiano,D,,
Anyalebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,
Balwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,,
Biswal,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,,
Bryant,N,, Buhay,C,, Burch,P,, Burrell,K,, Calderon,E,,
Cardenas,V,, Carter,K,, Cavazos,I,, Caesar,H,, Center,A,,
Chacko,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,,
Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Gouza,L,,
Davila,M.L., Davis,C,, Davy-Catroll,L,, De Anda,C,, Dederich,D,,
Delgado,O,, Denson,S,, Deramo,C,, Ding,Y,, Dinh,H,, Divya,K,,
Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,,
Egan,A,, Escotto,M,, Eugene,C,, Evans,C.A., Falls,T,, Fan,G,,
Fernandez,S,, Finley,M,, Flagg,N,, Forbes,L,, Foster,M,, Foster,P,,
Fraser,C.M., Gabisi,A,, Ganta,R,, Garcia,A,, Garner,T,, Garza,M,,
Gebregorgis,E,, Geer,K,, Gill,R,, Grady,M,, Guerra,W,, Guevara,W,,
Gunaratne,P,, Haaland,W,, Hamil,C,, Hamilton,C,, Hamilton,K,,
Harvey,Y,, Haylak,P,, Hawes,A,, Henderson,N,, Hernandez,J,,
Hernandez,R,, Hines,S,, Hladun,S.L., Hodgson,A,, Hognes,M,,
Hollins,B,, Howells,S,, Hulik,S,, Hume,J,, Idlebird,D,, Jackson,A,,
Jackson,L,, Jacob,L., Jiang,H,, Johnson,B,, Johnson,R,, Joliver,A,,
Karpathy,S,, Kelly,S,, Kelly,S,, Khan,Z,, King,L,, Kovar,C,,
Kows,C,, Kraft,C.L., Lebow,H,, Levay,J,, Lewis,L., Li,Z,, Liu,J,,
Liu,J., Liu,W,, Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshew,L., Louleaged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawliny,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morjan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwankemeloh,O., Okunonu,G., Olarnpunagoon,A., Pal,S., Parke,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plouffe,F., Poindexter,A., Popovic,D., Prims,B., Pu,L., Lu.,
Puafo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sander,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sned,A., Sodergren,B., Song,X.-Z., Sorrelle,R., Soza,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Williams,G., Willson,R., Wiczek,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 263730)
Worley,K.C.
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 263730)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24942637.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GAJ1
Center clone name: CH230-3G10

----- Summary Statistics
Assembly program: Atlas
Consensus quality: 211248 bases at least Q40
Consensus quality: 217135 bases at least Q30
Consensus quality: 220721 bases at least Q20
Estimated insert size: 226625; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 262005: contig of 262005 bp in length
* 262006 262105: gap of unknown length
* 262106 263730: contig of 1625 bp in length.

FEATURES
source
Location/Qualifiers
1..263730
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-3G10"
1..1550
/note="wgs_end_extension
clone_end:5p6"
728..1335
/note="clone boundary
clone_end:3p6"
misc_feature
misc_feature

```

```

site:ECORI
end sequence: BH309953"
misc_feature 31375..33614
  /note="wgs_contig"
misc_feature 40354..44099
  /note="wgs_contig"
misc_feature 123395..124915
  /note="wgs_contig"
misc_feature 179305..180899
  /note="wgs_contig"
misc_feature 210644..211731
  /note="wgs_contig"
misc_feature 211782..213121
  /note="wgs_contig"
misc_feature 246452..247376
  /note="clone_boundary"
  clone_end:17
site:ECORI
end sequence: BH309950"
misc_feature 249386..250543
  /note="wgs_end_extension"
  clone_end:T7"
misc_feature 260312..262005
  /note="wgs_end_extension"
  clone_end:T7"

ORIGIN
Query Match 85.0%; Score 17; DB 2; Length 263730;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTG 17
    |||||
Db 167351 GGTCCATCTCATGTTG 167335

RESULT 24
LOCUS MMXPAC6/c 294 bp DNA linear ROD 06-FEB-1997
DEFINITION M.musculus XPAC Xeroderma Pigmentosum group A Correcting gene, exon
6.
ACCESSION X74350
VERSION X74350.1 GI:440563
KEYWORDS Xeroderma Pigmentosum group A Correcting gene (XPAC); Xeroderma
Pigmentosum group A Correcting protein (XPAC).
SOURCE Mus musculus
  Mus musculus (house mouse)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 294)
  van Oostrom, C.T., de Vries, A., Verbeek, S.J., van Kreijl, C.F. and
  van Steeg, H.
  Cloning and characterization of the mouse XPAC gene
  Nucleic Acids Res. 22 (1), 11-14 (1994)
JOURNAL 94173654
PUBMED 8127648
REFERENCE 2 (bases 1 to 294)
  van Steeg, H.
  Direct Submision
  Submitted (27-JUL-1993) H. Van Steeg, National Inst of Public
  Health and, Environmental Protection, P O B 1, 3720 BA Bilthoven,
  NETHERLANDS
FEATURES
  source
    1..294
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="01a129"
      /db_xref="taxon:10090"
      /chromosome="4C2"
    39..2215
      /usedin=X74345:XPAC_CDS
      /label=ex6
    210..215
      polyA_signal

```

```

ORIGIN
Query Match 84.0%; Score 16.8; DB 10; Length 294;
Best Local Similarity 90.0%; Pred. No. 7.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
    |||||
Db 108 GGTCCATCTCATGTTGATG 89

RESULT 25
AF354052 576 bp mRNA linear ROD 02-MAR-2002
LOCUS AF354052/c
DEFINITION Mus musculus xeroderma pigmentosum complementation group A mRNA,
complete cds.
ACCESSION AF354052
VERSION AF354052.1 GI:19070698
KEYWORDS
SOURCE Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 576)
  Chen, H. and Huang, C.-H.
  A novel spliced form of XPA in the mouse kidney (Rhcg)
  Unpublished
REFERENCE 2 (bases 1 to 576)
  Chen, H. and Huang, C.-H.
  Direct Submission
  Submitted (27-FEB-2001) Biochemistry and Molecular Genetics, New
  York Blood Center, 310 East 67th Street, New York, NY 10021, USA
FEATURES
  source
    1..576
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
      /tissue_type="kidney"
    1..39
      40..285
      /note="XPA, alternatively spliced"
      /codon_start=1
      /product="xeroderma pigmentosum complementation group A"
      /protein_id="AAL83963.1"
      /db_xref="GI:19070698"
      /translation="MDSYLMNHFDLPDSCGRDADDKHLITTKAEQEVLLKDCDLE
      KRERLRFLVKGNFRHSQMGDMKYLKIQVSSGRAGS"
    555..560
    polyA_signal

ORIGIN
Query Match 84.0%; Score 16.8; DB 10; Length 576;
Best Local Similarity 90.0%; Pred. No. 7.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20
    |||||
Db 453 GGTCCATCTCATGTTGATG 434

RESULT 26
LOCUS MMXPAC7/c 938 bp mRNA linear ROD 17-MAY-1994
DEFINITION MMXPAC7
  M.musculus mRNA for XPAC Xeroderma Pigmentosum group A Correcting
  protein.
ACCESSION X74351
VERSION X74351.1 GI:440564
KEYWORDS Xeroderma Pigmentosum group A Correcting gene (XPAC); Xeroderma
Pigmentosum group A Correcting protein (XPAC).
SOURCE Mus musculus
  Mus musculus (house mouse)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 938)

```

```

AUTHORS      van Oostrom,C.T., de Vries,A., Verbeek,S.J., van Kreijl,C.F. and
              van Steeg,H.
TITLE         Cloning and characterization of the mouse xPAC gene
JOURNAL      Nucleic Acids Res. 22 (1), 11-14 (1994)
MEDLINE      94173654
REFERENCE     PUBMED
AUTHORS      van Steeg,H.
TITLE         Direct Submission
JOURNAL      Submitted (27-JUL-1993) H. Van Steeg, National Inst of Public
              Health and, Environmental Protection, P O B 1, 3720 BA Bilthoven,
              NETHERLANDS

FEATURES
source
location/Qualifiers
    1..938
       /organism="Mus musculus"
       /mol_type="mRNA"
       /strain="Ola 129"
       /db_xref="taxon:10090"
       /cell_line="embryonic stem cell line E14"
    1..938
       /gene="xPAC Xeroderma Pigmentosum group A Correcting gene"
    92..910
       /gene="XPAC Xeroderma Pigmentosum group A Correcting gene"
       /codon_start=1
       /product="XPAC protein"
       /protein_id="CA52393.1"
       /db_xref="GI:440565"
       /db_xref="GOA:Q64267"
       /translation="MATEEKQTSPKPAADPEAPDAVAASVERKRALMLRQAR
LAAPFPAATAGVASVKAPKMDITGGFTLEBEKHETGINVHEPGVMEDDTYI
ICEGCKEFMDSYLNNHFDLPDCSDPADDKHLITTEAKQSYLNDCELRKEPDYD
LRFPLKKNPRIHSOWGMKLVIKLGVKVALWGSOEALIDPAKEVRLQENREMKOKKKE
DKXVELRRRAIRSSVWKRETTHQHVEYGFEENLEDMDYRKTCILGHETYEKM"
    933..938
       /gene="XPAC Xeroderma Pigmentosum group A Correcting gene"

ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 938;
Best Local Similarity 90.0%; Pred. No. 7,8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 GGTCATCTCATGTGTGATG 20
          ||||||| ||||| |
Db      831 GGTCATATTCTATGTTGGTG 812

RESULT 27
LOCUS        AX566974               1029 bp    DNA                linear      PAT 29-NOV-2002
DEFINITION   Sequence 173 from Patent WO02077021.
ACCESSION    AX566974
VERSION      AX566974.1 GI:26001549
KEYWORDS
SOURCE
ORGANISM     Streptococcus pneumoniae
              Streptococcus pneumoniae
              Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
              Streptococcus.
REFERENCE
AUTHORS      Maeignani,V., Tettelin,H. and Fraser,C.
TITLE         Streptococcus pneumoniae proteins and nucleic acids
JOURNAL      Patent: WO 02077021-A 173 03-OCT-2002;
              Chiron Spa (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES
source
location/Qualifiers
    1..1029
       /organism="Streptococcus pneumoniae"
       /mol_type="unassigned DNA"
       /db_xref="taxon:1313"

ORIGIN
Query Match      84.0%; Score 16.8; DB 6; Length 1029;
Best Local Similarity 90.0%; Pred. No. 7,8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      1  GGTGCATCATCATGTTGATG 20
        ||||| ||||| |||||
Db      975  GGTGCATCATCATGTTATG 994

RESULT 28
LOCUS   AKI15474               2138 bp    mRNA    linear   INV 30-NOV-2002
DEFINITION   Clona intestinalis cDNA, clone:cleg055b24, full insert sequence.
VERSION      AKI15474.1  GI:23587571
KEYWORDS     FLI cDNA.
SOURCE       Clona intestinalis
ORGANISM     Clona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cloniidae; Clona.

REFERENCE
AUTHORS   1 Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
            Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasakura,Y.,
            Nakayama,A., Ishikawa,H., Inaba,K. and Satoh,N.
            A cDNA resource from the basal chordate Clona intestinalis
            Genesis 33 (4), 153-154 (2002)
JOURNAL   22191024
MEDLINE   12203911
PUBMED    2 (bases 1 to 2138)
REFERENCE 2 Satou,Y. and Satoh,N.
            Direct Submission
            Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
            Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            (E-mail:satoh@ascidiam.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,
            Fax:81-75-705-1113)
COMMENT   Clona intestinalis cDNA Project (URL:
            http://ghost.zool.kyoto-u.ac.jp/index1.html).
            Location/Qualifiers
FEATURES             1..2138
                     /organism="Clona intestinalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:7719"
                     /clone="cleg055b24"

ORIGIN
Query Match      84.0%; Score 16.8; DB 3; Length 2138;
Best Local Similarity 90.0%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy      1  GGTGCATCATCATGTTGATG 20
        ||||| ||||| |||||
Db      1185  GGTGCATCATCATGTTATG 1166

RESULT 29
LOCUS   AE007333/c             11344 bp    DNA    linear   BCT 31-AUG-2001
DEFINITION   Streptococcus pneumoniae TIGR4 section 16 of 194 of the complete
            genome.
VERSION      AE007333 AE005672
KEYWORDS
SOURCE       Streptococcus pneumoniae TIGR4
            Streptococcus pneumoniae TIGR4
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1 (bases 1 to 11344)
            Tettelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D.,
            Peterson,S., Heidelberg,J., Deboy,R.T., Haft,D.H., Dodson,R.J.,
            Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D.,
            Umayam,L.A., White,O., Salzberg,S.L., Lewis,M.R., Radune,D.,
            Holtzapple,E., Khouri,H., Wolf,A.M., Ueberback,T.R., Hansen,C.L.,
            McDonald,L.A., Feldblyum,T.V., Angiuoli,S., Dickinson,T.,
            Hickey,E.K., Holt,I.E., Loftus,B.J., Yang,F., Smith,H.O.,
            Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and
            Fraser,C.C.

```

TITLE Complete genome sequence of a virulent isolate of Streptococcus pneumoniae  
JOURNAL Science 293 (5529), 498-506 (2001)  
MEDLINE 21357209  
PUBMED 11463916  
REFERENCE 2 (bases 1 to 13344)  
AUTHORS Tetteelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., Deboy,R.T., Haft,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Lewis,M.R., Radune,D., Holtzapple,E., Knouri,H., Wolf,A.M., Uetebach,T.R., Hansen,C.J., McDonald,L.A., Felblyum,T.V., Angluoli,S., Gesuwan,P., Hickey,E.K., Holt,I.E., Loftus,B.J., Uwal,M.L., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.  
TITLE Direct Submission  
JOURNAL Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
FEATURES  
source 1. .11344  
/organism="Streptococcus pneumoniae TIGR4"  
/mol\_type="genomic DNA"  
/strain="TIGR4"  
/db\_xref="taxon:170187"  
90. .194  
/gene="SP0172"  
90. .194  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK74353.1"  
/db\_xref="GI:14971639"  
/translation="MFIETITLLPLADPYVEGLSGYVELFCSP"  
255. .2204  
/gene="SP0173"  
255. .2204  
/gene="SP0173"  
/codon\_start=1  
/transl\_table=11  
/product="DNA mismatch repair protein HexB"  
/protein\_id="AAK74354.1"  
/db\_xref="GI:14971640"  
/translation="MSHITLPEMLANOIAGSEVIERPAAVYKELVENADAGSQIITIEERGLKKVOITONGHIADEVLAIRRHATSKIQADLPFRITLPGFGLPSIASVSUTLLTAVDASHGTXKYARGEVEEYIPATSPVGTQCYVEDLPFNPAIKYMSQQAELSHIIDIVNRIGLAHPEISFSLISDGKENTRTAGTQLRQAIAGIYGLVSAKMIIEISNDLDFEISGFVSLPELTPANRNYISLFLNGRYIKNPLANRAIDGFGSKLMGRPELAVIHIHIDPYLADVNVHPTKQEVRISSKEKELMTLVSEAIANSLEKQTLIPD ALENLAKSTVNRBEKVEOTILPLKENTLYEKTEPERSQTEVADVOVLTDEGDLT LAKEITLDELTYPAKLAFAERKPRANDOLDHPELDIASIDKAVYDKLEREASFPBELFEGQNHGTYLFAQGRDGLYTIIDQHAQOEKVEKREYESIGNDOSQQLIVYIIEFPADDAALKEKRMPLIEVGVFLAEYGNQFLIEHPLIWMABEIESGTYEMCDMLLTKEVSIKRYRAELAIMMSCRSIKANRDIHDSARQLLYOLSDCNPNYCPHGRPVLVHFTKSDMEKRRRIQENTHSLRELKGY"  
complement (2340. .2432)  
/gene="SP0174"  
complement (2340. .2432)  
/gene="SP0174"  
/note="identified by Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAK74355.1"  
/db\_xref="GI:14971641"  
/translation="MVKHNFVDVDTGKISKHCFEITDKTDV"  
complement (2529. .2996)  
/gene="SP0175"  
complement (2529. .2996)  
/gene="SP0175"  
/note="identified by match to PFAM protein family HM PF00885"  
/codon\_start=1

gene  
CDS  
/transl\_table=11  
/product="riboflavin synthase, beta subunit"  
/protein\_id="AAK74356.1"  
/db\_xref="GI:14971642"  
/translation="MNTYEGNIVANNIKIGIVARFNEFITSKLISGALDNLKRENVN EKDIEVAVWGAPEIPIIASAKMAKSKYDAILCLGAVIRGNTSHYDVCEVSGKIQ ISLNSLIPVMEGVLTDTIIOALERAGTKAKNGSECAQATEMNLIRTLDA"  
complement (2997. .4202)  
/gene="SP0176"  
complement (2997. .4202)  
/gene="SP0176"  
/note="identified by match to PFAM protein family HM PF00926"  
/codon\_start=1  
/transl\_table=11  
/product="3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II"  
/protein\_id="AAK74357.1"  
/db\_xref="GI:14971643"  
/translation="MEYRKIOEALQKRLVVIDDKRENEGDLICGAQAATTEN VNFATYAKGLICMPSESLANQMLSPVVENNTDNHKTAFVSIIDYKETTGISAE RGLTARMGVADIEITPSDPRRGHMFPLIAKKGVLERNGTEATVDLLKLGLKECL CCEIIMNDGKMRPTDILIOSKKNITPLITTKLOERYKRYDOLVERVSTYVMPRTYGNFAISYTDKUNGEHHALINGNIEDBANVLCRHSECTIGDYLSRCOCGOFDRA MKMIVENGSGVLLYLRQEGRTGLINKAKVHLDDQGMQDLDANLAFGBDLREYH GAKMLKDQLQSHTLTNNPDKXVQLERKGYTISRSISIEIANPYDSFYLETXKRM GHILNMEK"  
complement (4222. .4857)  
/gene="SP0177"  
complement (4222. .4857)  
/gene="SP0177"  
/note="identified by match to TIGR protein family HM TIGR00187"  
/codon\_start=1  
/transl\_table=11  
/product="riboflavin synthase, alpha subunit"  
/protein\_id="AAK74358.1"  
/db\_xref="GI:14971644"  
/translation="MFTGIEIEIGKVERIOKDSHNCKLSIKASKILTDIHLGSIANV GCLTVHFHQSFVDMVNETWRTALTLKHSEVNLSEALSVNGLGHHVVTGHI DGTGKISIKKDDNNAVYQNTQKEIIDLIVERKSTIIDIGSLTAVAKSVNFSVSI PHTLEQTLISKQVGYVNLENDILGKTVQKLMNSKSELSKELLQNGF"  
complement (4842. .5942)  
/gene="SP0178"  
complement (4842. .5942)  
/gene="SP0178"  
/note="identified by match to TIGR protein family HM TIGR00326"  
/codon\_start=1  
/transl\_table=11  
/product="riboflavin biosynthesis protein RibD"  
/protein\_id="AAK74359.1"  
/db\_xref="GI:14971645"  
/translation="MSDSKIKKLAIKLAQKAGVYVNPVGVAYIKVDNHIIGCYHE FFGGHAERNALKNCRSPVAGTLYVLPCHRGKTPCIDAIIDSGITKRVVIGSLD CNPIYSGKGVKILENNILOVYGLIENECMLIKSFRYITQHPVYPMKXAMSGDI IATKNOGKWTIEBARQVHQLRHVSAIMVGVYIODDPLTICRLEBGNP IRII CDTHLRPTLSKIVKTNADIKTYIATSSEDGKNKXKIVONHGCETLSIKKKNHIDLS LMQHGNGQISLVLEGSIMMWSALBQIYDELKITYAPRVFGSNAKFPVGGGISTL PNDATRLRPVAFSQGNDYLIIEVYIPCSQ"  
6347. .6940  
/gene="SP0179"  
6347. .6940  
/gene="SP0179"  
/note="identified by match to PFAM protein family HM PF01330"  
/codon\_start=1  
/transl\_table=11  
/product="Holliday junction DNA helicase RuvA"  
/protein\_id="AAK74360.1"  
/db\_xref="GI:14971646"  
/translation="MYAYLKGIITKITAKYIVLETNGIGIYLVHVPYAYSGQVQGEA OIVYHQVVRDEAHLLYGFRSEDEKTLFSLISVSGIGPVSALAIITAADNGLVOAIE

gene  
CDS  
TNTNLYLKPRPRIGKTKAQOMVLDIEGGVVAVAGDDLPARKAVQASAEINQELBEAMEN  
LALGYKATLEKLIKIKKFFEGSTDTLANTYKSKALNKLTVK"  
6977..7513  
/gene="SP0180"  
6977..7513  
/gene="SP0180"  
/codon\_start=1  
/transl\_table=11  
/product="DNA-3'-methyladenine glycosylase I"  
/protein\_id="AAK74361.1"  
/db\_xref="GI:14971647"  
/translation="MTNPLYTAYHDEMGQPLHDDOVLPELLCMETVQAQSLWETV  
LNKRAAREVHSGYQIHSVAMETDTEALLENPAIIIRRAKLPATRAAAQFIQADAY  
GSPFDLYMSPEFGKTVVNDVPDYKQAPAKTPLSEKLAKDLKKQKFTGPVAVLSTFQ  
AAGLYDDHMSQENCKGLK"  
7510..8187  
/gene="SP0181"  
7510..8187

gene	complement (8324 . .9355)
	/gene="SP0182"
CDS	complement (8324 . .9355)
	/gene="SP0182"

Query Match	84.0%;	Score 16.8;	DB 1;	Length 11344;
Best Local Similarity	90.0%;	Pred. No. 9e+02;		
Matches 18; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY            1 GGTCATACTCATGTGATG 20  
             |||||  
Db           8381 GGTCCATGCTCATGTAGATG 8362

RESULT 30			
AR218808/c			
LOCUS	AR218808	14273 bp	DNA
DEFINITION	Sequence	40 from patent US 6420135.	linear
ACCESSION	AR218808		
VERSION	AR218808.1	GI:23319742	

REFERENCE 1. (base 1 to 14273)  
AUTHORS Kunsch, C.A., Choi, G.H., Dillon, P. S., Rosen, C.A., Barash, S.C.,  
Ramon, M.R. and Dougherty, B.A.  
TITLE Streptococcus pneumoniae polysaccharide and sequences  
JOURNAL Patent : US 6420133-A 40 16-Jun-2002;  
FEATURES Location/Qualifiers  
source 1. .14273

ORIGIN

Query Match	84.0%	Score 16.8;	DB 6;	Length 14273;
Best Local Similarity	90.0%	Pred. No. 9.1e+02;		
Matches 18; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

```

Oy      1  GGTCCATACTCATGTGTATG  20
          |||||
Db      13385 GGTCCATGCTCATGTAGATG  13366

```

RESULT	31
BD003720/c	
LOCUS	BD003720 14273 bp DNA linear PAT 31-JAN-2002
DEFINITION	Polymnucleotide of Streptococcus pneumoniae and sequence.
ACCESSION	BD003720
VERSION	BD003720.1 GI:18631681
KEYWORDS	JP 2001501833-A/40.
SOURCE	unidentified
ORGANISM	unclassified
REFERENCE	1 (bases 1 to 14273)
AUTHORS	Kumesh,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Bara,S.C., Fannon,M. and Dougherty,B.A. Polynucleotide of Streptococcus pneumoniae and sequence Patent: JP 2001501833-A 40 13-FEB-2001;
TITLE	HUMAN GENOME SCIENCES INC
JOURNAL	OS Unidentified
COMMENT	

```

PI CHARLES A KINGSCH, GIL H CHOI, PATRICK J DILLON, CRAIG A ROSEN, PI
STEVEN C BARASH
PI MICHAEL FANNON, BRIAN A DOUGHERTY
PC C12N15/09, A01K67/027, C07K14/315, C07K16/12, C12N1/15, C12N1/19,
PC C12N1/21,
PC C12N5/10, C12P21/02, C12Q1/68, G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC Strandedness: Double;
CC Topology: Linear;
PH key Location/Qualifiers
FT source 1..14273
FT /organism='Unidentified',

```

```
FEATURES
source      Location/Qualifiers
1. .14273
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
```

Query Match	84.0%	Score	16.8	DB 62	Length	14273			
Best Local Similarity	90.0%	Pred.	No.9.1e+02						
Matches	18	Conservative	0	Mismatches	2	Indels	0	Gaps	0

```

QY      1  GGTCCATACTCATGTGATG  20
          |||||
Db      13385 GGTCCATGCTCATGTAGATG  13366

```

RESULT	32
HSJ98A12	
LOCUS	HSJ98A12
DEFINITION	Human DNA sequence from clone RPJ-398A12 on chromosome 6p23-25.1,
ACCESSION	U02113.2
VERSION	AL021332.1 GI:2804159
KEYWORDS	HMG.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens (human)

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
(pages 1 to 90169)	Williams,S.	Direct Submission	Submitted (04-MAR-2003)	Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hummery@sanger.ac.uk
				Clone requests: clonerequest@sanger.ac.uk
				Genome Center

Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

-----  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>

RP3-398A12 is from the library RPCT-3 constructed by the group of Plietier de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCYPAC2.

#### FEATURES

source  
 1..50169  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="RZPD:RPCTP704A12398"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /map="p23-25.1"  
 /clone="RP3-398A12"  
 /clone\_1ib="RPCT-3"

#### ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 90169;  
 Best Local Similarity 90.0%; Pred. No. 1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
 |||||  
 DB 60420 GCTCCATCTCATGTTGATG 60439

RESULT 33  
 AC095009\_1/c  
 WPCOMMENT  
 Sequence split into 4 fragments LOCUS AC095009 Accession AC095009

Fragment Name	Begin	End
AC095009_0	1	110000
AC095009_1	100001	210000
AC095009_2	200001	310000
AC095009_3	300001	395260

Continuation (2 of 4) of AC095009 from base 100001 (AC095009 Rattus norvegicus clone CH2)

Query Match 84.0%; Score 16.8; DB 2; Length 110000;  
 Best Local Similarity 90.0%; Pred. No. 1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
 |||||  
 DB 94388 GGCCTATCTCATGTTGATG 94369

RESULT 34  
 BX537312/c  
 LOCUS  
 DEFINITION Danio rerio clone DKEX-6P9, WORKING DRAFT SEQUENCE, 4 unordered

ACCESSION BX537312  
 VERSION BX537312.4 GI:35209488  
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Danio rerio (zebrafish)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 McLaren S, Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthys; Cypriniformes; Cyprinidae; Danio. 1 (bases 1 to 121113)

COMMENT  
 Direct Submission  
 Submitted (23-Sep-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk On Sep 24, 2003 this sequence version replaced gi:33414383.

----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: zfish-help@sanger.ac.uk  
 ----- Project Information  
 Center project name: zK6P9

----- Summary Statistics  
 Assembly program: XGAP; version 4.5  
 Chemistry: Dye-terminator; 100% of reads  
 Consensus quality: 120006 bases at least Q40  
 Consensus quality: 120426 bases at least Q30  
 Consensus quality: 120458 bases at least Q20  
 Insert size: 120813; sum-of-contigs  
 Insert size: 126278; 4.2% error; agarose-fp  
 Quality coverage: 11.33x in Q20 bases; sum-of-contigs Quality coverage: 10.84x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 63519: contig of 63519 bp in length  
 \* 63520 63619: gap of 100 bp  
 \* 63620 88215: contig of 24596 bp in length  
 \* 88216 88315: gap of 100 bp  
 \* 88316 110692: contig of 22377 bp in length  
 \* 110693 110792: gap of 100 bp  
 \* 110793 121113: contig of 10321 bp in length.

FEATURES  
 source  
 1..121113  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEX-6P9"  
 /clone\_1ib="DanioKey"

misc\_feature  
 1..63519  
 /note="assembly\_fragment:00731  
 fragment\_chain:1"  
 misc\_feature  
 63620..88215  
 /note="assembly\_fragment:01659  
 fragment\_chain:1"  
 misc\_feature  
 88316..110692  
 /note="assembly\_fragment:00742.0"  
 110793..121113  
 /note="assembly\_fragment:00871"

#### ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 121113;  
 Best Local Similarity 90.0%; Pred. No. 1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20



[illegible]

Contact: humquery@sanger.ac.uk

-----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YVC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrx>

RP3-326L13 is from the library RP3-3 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

Location/Qualifiers

1. 127247

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="RZPD:RP3-326L13326"

/db\_xref="taxon:9606"

/chromosome="X"

/map="q21.1"

/clone="RP3-326L13"

/clone\_1fb="RP3-3"

-----

Query Match 84.0%; Score 16.8; DB 9; Length 127247;  
Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;  
Matches 18; Conservative 0; Mismatches 2;

Db 93812 GTCCATTCATCATGTGATG 93793

Oy 1 GGTCATCATCATGTGATG 20  
|||||  
|

RESULT 37

AP003886 129300 bp DNA linear PLN 21-AUG-2003

LOCUS Oyza sativa (japonica cultivar-group) genomic DNA, chromosome 8, BAC clone:OU1163\_G08, complete sequence.

ACCESSION AP003886

VERSION AP003886

KEYWORDS HG.

SOURCE Oyza sativa (japonica cultivar-group)

ORGANISM Oyza sativa (japonica cultivar-group)

REFERENCE 1. Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Oryza sativa nipponbare (GAS) genomic DNA, chromosome 8, BAC clone:OU1163\_G08  
Published Only in Database (2001)

2. (bases 1 to 129300)  
Sasaki, T., Matsumoto, T. and Yamamoto, K.  
Direct Submission  
Submitted (09-JUL-2001) Takuji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program, Kannondai



```

repeat_region      /rpt_family="L1"
                    6726. .6749
repeat_region      /rpt_family="AT_rich"
                    6981. .7031
repeat_region      /rpt_family="MER103"
                    7661. .7691
repeat_region      /rpt_family="(T)n"
                    8084. .8118
repeat_region      /rpt_family="AT_rich"
                    8449. .8642
repeat_region      /rpt_family="MIR"
                    8659. .8709
repeat_region      /rpt_family="(CAT)n"
                    8717. .9264
repeat_region      /rpt_family="L1"
                    9343. .9406
repeat_region      /rpt_family="AT_rich"
                    9778. .10068
repeat_region      /rpt_family="Alu"
                    10072. .10128
repeat_region      /rpt_family="L2"
                    10190. .10217
repeat_region      /rpt_family="AT_rich"
                    10360. .10560
repeat_region      /rpt_family="L1"
                    11312. .11379
repeat_region      /rpt_family="(TG)n"
                    11335. .12092
repeat_region      /rpt_family="L1"
                    12524. .12548
repeat_region      /rpt_family="AT_rich"
                    13448. .13780
repeat_region      /rpt_family="L1"
                    14263. .14289
repeat_region      /rpt_family="(A)n"
                    15057. .15212
repeat_region      /rpt_family="MIR"
                    15481. .15741
repeat_region      /rpt_family="Alu"
                    15820. .15844
repeat_region      /rpt_family="AT_rich"
                    16142. .16424
repeat_region      /rpt_family="L1"
                    16425. .16817
repeat_region      /rpt_family="L1"
                    16830. .17069
repeat_region      /rpt_family="L1"
                    17070. .17369
repeat_region      /rpt_family="Alu"
                    17370. .17610
repeat_region      /rpt_family="L1"
                    17800. .18064
repeat_region      /rpt_family="L1"
                    18485. .18505
repeat_region      /rpt_family="AT_rich"
                    19665. .19938
repeat_region      /rpt_family="L2"
                    20337. .20661
repeat_region      /rpt_family="MER1_type"
                    20779. .20822
repeat_region      /rpt_family="MALR"
                    20823. .21210
repeat_region      /rpt_family="MALR"
                    21436. .21743
repeat_region      /rpt_family="Alu"
                    21841. .22009
repeat_region      /rpt_family="L1"
                    22014. .22370
repeat_region      /rpt_family="MALR"
                    22609. .22827
repeat_region      /rpt_family="L1"
                    24733. .24754
repeat_region      /rpt_family="AT_rich"

```

```

repeat_region      25752. .25966
                    /rpt_family="L1"
repeat_region      26145. .26179
                    /rpt_family="AT_rich"
repeat_region      26183. .26454
                    /rpt_family="Alu"
repeat_region      26540. .26801
                    /rpt_family="MALR"
repeat_region      26830. .26989

Query Match      84.0%; Score 16.8; DB 9; Length 148416;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GGTGATGCTGATGTTGATG 20
        ||||| ||||| |||||
Db      57746 GGTGATGCTGCTGTTGATG 57765

RESUT 39
AC011144
LOCUS      Homo sapiens clone RP11-1024, WORKING DRAFT SEQUENCE, 5 unordered
           pieces.
ACCESSION  AC011144
VERSION     AC011144.4 GI:13357401
KEYWORDS    HTG; HTGS; PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 152635)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome, clone RP11-1024
Unpublished
2 (bases 1 to 152635)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckery,R., Boguslavskiy,L., Boukagalter,B.,
Brown,A., Casale,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lhoczeky,J., Lieu,C., Locke,K., Macdonald,P., Margulis,N.,
McSwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 16, 2001 this sequence version replaced gi:12583837.
All repeats were identified using RepeatMasker.
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIPR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2517
Center clone name: 1 O 24
----- Summary Statistics
Sequencing vector: M13; M77815; 3% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 151169 bases at least Q40
Consensus quality: 151789 bases at least Q30
Consensus quality: 152081 bases at least Q20

```

Insert size: 170000; agarose-fp  
 Insert size: 152235; num-of-contigs  
 Quality coverage: 6.2 in Q20 bases; agarose-fp  
 Quality coverage: 6.9 in Q20 bases.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 21727: contig of 21727 bp in length  
 \* 21728 21827: gap of 100 bp  
 \* 21828 69977: contig of 48150 bp in length  
 \* 69978 70077: gap of 100 bp  
 \* 70078 111601: contig of 41524 bp in length  
 \* 111602 111702: gap of 100 bp  
 \* 111702 146223: contig of 34522 bp in length  
 \* 146224 146324: gap of 100 bp  
 \* 146324 152635: contig of 6312 bp in length.

FEATURES  
 source  
 1. 152635  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="RP11-1024"  
 /clone\_1lb="RPC1-11 Human Male BAC"  
 1. 21727  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left"  
 misc\_feature  
 21828..69977  
 /note="assembly\_fragment"  
 misc\_feature  
 70078..111601  
 /note="assembly\_fragment"  
 misc\_feature  
 111702..146223  
 /note="assembly\_fragment"  
 misc\_feature  
 146324..152635  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right"

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 2; Length 152635;  
 Best Local Similarity 90.0%; Pred.No.1e+03; Mismatches 2; Indels 0; Gaps 0;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 GGTCCATCTCATGTTGATG 20  
 Db 7397 GGTCCATGCTCTTGTGATG 7416

RESULT 40  
 AC146044/c 155646 bp DNA linear PRI 18-DEC-2003  
 LOCUS AC146044  
 DEFINITION Pan troglodytes BAC clone RP43-109N23 from 7, complete sequence.  
 ACCESSION AC146044  
 VERSION AC146044.2 GI:38424257  
 KEYWORDS HTG.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
 TITLE The sequence of Pan troglodytes BAC clone RP43-109N23  
 JOURNAL Unpublished (2001)  
 REFERENCE 2 (bases 1 to 155646)  
 AUTHORS Wilson,R.  
 JOURNAL Sequencing of Pan troglodytes  
 TITLE Unpublished (2001)  
 REFERENCE 3 (bases 1 to 155646)

AUTHORS  
 TITLE Wilson,R.K.  
 JOURNAL Direct Submission  
 Submitted (01-NOV-2003) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 4 (bases 1 to 155646)  
 REFERENCE Wilson,R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-2003) Genetics, Genome Sequencing Center, 4444  
 Forest Park Parkway, St. Louis, MO 63108, USA  
 5 (bases 1 to 155646)  
 REFERENCE Wilson,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-DEC-2003) Department of Genetics, Washington  
 University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Nov 19, 2003 this sequence version replaced gi:33386996.  
 COMMENT ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: submissions@watson.wustl.edu  
 ----- Summary Statistics  
 -----  
 Center project name: C\_PRI109N23  
 -----

NOTICE: This sequence may not represent the entire insert of this  
 clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. Wes Warren,  
 Department of Genetics, Washington University, St. Louis MO. For  
 additional information about the map position of this sequence, see  
 http://genome.wustl.edu

SOURCE INFORMATION:  
 The RPC1-43 BAC library has been constructed by Chung-Li Shu. DNA  
 was isolated from white blood cells obtained from a male chimpanzee  
 (Pan troglodytes, 'Cint', Yerkes #C0477; birthdate: 6-6-80). The  
 clone and detailed information can be obtained from Regen  
 (http://www.regen.com) or Pieter de Jong and co-workers at  
 http://www.bacpac.chori.org.

NEIGHBORING SEQUENCE INFORMATION:  
 This sequence is the entire insert of the clone.  
 Location/Qualifiers  
 1. 155646  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /chromosome="7"  
 /map="7"  
 /clone="RP43-109N23"  
 /clone\_1lb="RPC1-43"

FEATURES  
 source  
 1. 155646

ORIGIN  
 Query Match 84.0%; Score 16.8; DB 9; Length 155646;  
 Best Local Similarity 90.0%; Pred.No.1e+03; Mismatches 2; Indels 0; Gaps 0;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 GGTCCATCTCATGTTGATG 20  
 Db 106829 GGTCCATGCTCTTGTGATG 106810

[illegible]

FEATURES	source
gene	<p>1. 164572</p> <p>/organism="Oryza sativa (japonica cultivar-group)"</p> <p>/mol_type="genomic DNA"</p> <p>/variety="Nipponbare"</p> <p>/sub_species="japonica"</p> <p>/db_xref="taxon:39947"</p> <p>/chromosome="4"</p> <p>/clone="OSJNBa0088122"</p> <p>/clone_1lb="CUGT-OSJNBa"</p> <p>/complement (3174. .3644)</p> <p>/gene="OSJNBa0088122.1"</p> <p>/complement (3174. .3644)</p> <p>/gene="OSJNBa0088122.1"</p> <p>/codon_start=1</p> <p>/protein_id="CAD1569.2"</p> <p>/db_xref="GI:38567905"</p> <p>/translation="MAAVADVAAALRGPDALNPPDLAASRTAPPASSADGIRAA AAAAAALSDHRAAGCGIAPAAAGSAHQRGSSAAATTTGSGAOGEGSGAAGS QYFLDEALFETPQFLRNMAAGMMFPRLSPSSSDSPPSSEAGSLMSTRD"</p> <p>7871. .8788</p> <p>/gene="OSJNBa0088122.2"</p> <p>/join(7871. .8396,8445. .8788)</p> <p>/gene="OSJNBa0088122.2"</p> <p>/codon_start=1</p> <p>/protein_id="CAD1570.2"</p> <p>/db_xref="GI:38567906"</p> <p>/db_xref="SPTREMBL:Q7XU51"</p> <p>/translation="MRREGEKGGVFSALSGKDGEADGKLTGDERRRRRRMFDKED GRKAPLAGLRGKGCRRRTLRDGAALRPSGAMRRTLRGSEAGSEMGCGRGTG WRGCVLGRGGAAGDEATLTTHPEHGKMTIDRAGRGVTRKXVLTGMRSAVRRFRGR RENAADGALDGGSGGCGPPRCRCGVSGSASAEHRKPRGGCAPLPGRKKEKJAVRRAR PFPFWERGRHEBLCLALDARGVRARDDAGDDCKDCKDCLLEASTDTRMVSEQAT TEGDVSHPEIPR"</p> <p>/complement (26784. .28809)</p> <p>/gene="OSJNBa0088122.3"</p> <p>/complement (join(26784. .27108,27643. .27740,28624. .28809)</p> <p>/gene="OSJNBa0088122.3"</p> <p>/codon_start=1</p> <p>/protein_id="CAD1571.2"</p>
gene	<p>12447439</p> <p>Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Wang, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K., Zhou, B., Chen, Z. H., Hao, P., Zhang, L., Mu, M., Zhang, R. O., Guan, J. P., Fu, G., Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F., Tu, Y. F., Jia, J., Yin, H. P., Zhang, Y., Cai, Z., Chen, D., Kang, H., Chen, X. Y., Shao, C. Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W., Wang, L. J., Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L., Zhu, F. H. and Hong, G. F.</p> <p>Direct Submission</p> <p>Submitted (26-SEP-2001) Han Bin, National Center for Gene Research, Chinese Academy of Sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests: bhan@ncgr.ac.cn</p> <p>Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSJNBa0088122.</p> <p>On Nov 30, 2003 this sequence version replaced gi:32492239.</p> <p>Web site: <a href="http://www.ncgr.ac.cn">http://www.ncgr.ac.cn</a></p> <p>----- Summary Statistics -----</p> <p>Assembly program: phrap</p> <p>This is a complete sequence.</p> <p>Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (<a href="http://www.softberry.com/">http://www.softberry.com/</a>), genescan (<a href="http://CCR-081.mit.edu/GENSCAN.html">http://CCR-081.mit.edu/GENSCAN.html</a>), tRNAscanMM (<a href="http://genemark.biology.gatech.edu/genemark/">http://genemark.biology.gatech.edu/genemark/</a>), tRNAscan-SB (Sean Eddy, <a href="http://genome.wustl.edu/eddy/tRNAscan-SB/">http://genome.wustl.edu/eddy/tRNAscan-SB/</a>), searches of the complete sequence against NCBI non-redundant protein database (nr) (<a href="http://nbl.nlm.nih.gov/blast/db">http://nbl.nlm.nih.gov/blast/db</a>) and the EST database at NCGR. Location/Qualifiers</p>
REFERENCE	
AUTHORS	
PUBMED	

gene  
CDS  
/db\_xref="GI:38567907"  
/db\_xref="SPTREMBL:Q7XU50"  
/translation="MISGLINVSALGFGMTAAVAFCARFVCCARRADASARP  
HSPVDFADPSPDFDRPIEHSRSGLEPLAAVAIPMKYCNCAFSBEDTQSCISSE  
YKEDILRIVPICHNFHLYCDLAWLKQTCIPICRISKELPDGKSTVSAFPTMSQP  
PILPSSVNPSTSHPLPVHOEHRSHODGPDMPSEVVEIQR"  
30110..30921  
/gene="OSJNBA0088122.4"  
join(30110..30283,30363..30683,30769..30921)  
/gene="OSJNBA0088122.4"  
/codon\_start=1  
/protein\_id="CAD41572.3"  
/db\_xref="GI:38567908"  
/db\_xref="SPTREMBL:Q7XU49"  
/translation="MARRGVDDPMWRSDPCVYLTPMCORPLYCGDRCOLASRNP  
NIRGRFPFCYRETRITACAYIEWDTEPNVLDLTTCLOEGRAYFASSTEOYLQK  
AAVEROCOCSDMRVLTALPMEARPCRGDCQVLRSGINPTLGRFPVYNIL  
DDPMHPRRCYREMITRRLVLTTPSRVQLELPEQRYVTARFERGBGSSRG"  
complement(43323..45005)  
/gene="OSJNBA0088122.5"  
complement(43323..45005)  
/gene="OSJNBA0088122.5"  
/codon\_start=1  
/protein\_id="CAD41573.3"  
/db\_xref="GI:38567909"  
/db\_xref="SPTREMBL:Q7XU48"  
/translation="MTTCADQDTGCAFFAPLLSSKGAENVILVADDEAEBOQPAVL  
SKPGRLLAAVEASVSLGAFVTPSPMFTCSARGEARSLGLAFPMILITLILYL  
SMISMLFGLHGLAGSLAIGFANITGVSLGLAMGEPICGOAFGAGNVALLG  
HVRITVLTQSGINLPLVTCAGLAIAIHLPINVLVYVIGLAKVALASLVANLIVL  
FLAIVFLKGVHRTGFLLSAESFRGWELSLPSCVSCLEMYEIMILICGL  
LINPQATVASMGLIQTSLIYIPSSLSFGVSTVSNELGQDEESRAATVGLV  
GFCFAPASAFALVRNWSMFTADPAIVALTAVLPLIGICEHGENCPQTCGVLR  
GSARPKDASINLRSFVLVGTVALVMAWFLDPRGLMFLGAQAQCTVYMLLVIG  
RTDMAEAKRSQLTGAGANMSEDDRAADKSRLLPVDTVERSDHTDR"  
60927..63084  
/gene="OSJNBA0088122.6"  
join(60927..61265,62629..63084)  
/gene="OSJNBA0088122.6"  
/codon\_start=1  
/protein\_id="CAD41574.3"  
/db\_xref="GI:38567910"  
/db\_xref="SPTREMBL:Q7XU47"  
/translation="MDEIGPSPRGAEAGDLTSTAREAGOROGAGVRSQOPASPV  
YOOPATNGINGMACEVPGPHIIVFORLOHAKAIAPICIAAKNAFVHTSSYKTR  
LILISTSEBOLQILPODROBCKROILSLYRALMEMWMLCGTTAVGPEPELSS  
SSSINOSRGPFLDPTQAVTSTYLMERSRSRSDTSSYVANKSDASTOREVMM  
LOGRRMTVTITVRKMGFTMAPLITTSKTYISTNFRGKVDIVL"  
complement(63664..66067)  
/gene="OSJNBA0088122.7"  
complement(63664..66067)  
/gene="OSJNBA0088122.7"  
/codon\_start=1  
/protein\_id="CAD41575.3"  
/db\_xref="GI:38567911"  
/db\_xref="SPTREMBL:Q7XU46"  
/translation="MPESCVAALVLYLDPEICKARLBARFGAASADCVAAKLPA  
NYRVALAALADDSDGSGATGNGSRSSAAMIKKEIVATLCKPTPEDGKTFEM  
EKNGKICISISKMAITGIDDRYKMSLTSEBSRPHVATLQOIMLEVAGEIDFC  
EPAGSYSLFRLQGLGPHKRYGRVYGVSHIGMINKPRFPLSTSDDOQATSOYLLN  
EPGNMILYHVGDFVSSDQLTNLKFSMWQIDCTHKGGLCVDSVFYIPKGRHBDCT  
ICK"  
complement(70237..74263)  
/gene="OSJNBA0088122.8"  
complement(70237..74263)  
/gene="OSJNBA0088122.8"  
/codon\_start=1  
/protein\_id="CAD41576.3"  
/db\_xref="GI:38567912"  
/db\_xref="SPTREMBL:Q7XU45"  
/translation="MKKKTLANVSGLPSARMROSBRASWVYLSPTCAVACTSLVS

gene  
CDS  
/db\_xref="GI:38567913"  
/db\_xref="SPTREMBL:Q7XU44"  
/translation="MSAPSSPHSSSLFSLHLIRTSYARATGNGRESGRROY  
MRRRRSYASPARPPSPGSRWATMAALSLEGAALAAEAERRRVYSGSQRLARGA  
YVDNPPPTKODSNIHSEVNGCSVLCRAMKTNLANCASXPYSEQFPGTPY"  
80941..86909  
/gene="OSJNBA0088122.10"  
join(80941..82090,82295..82536,82597..84075,84346..84484,  
86395..86640,86692..86909)  
/gene="OSJNBA0088122.10"  
/codon\_start=1  
/protein\_id="CAD41578.3"  
/db\_xref="GI:38567914"  
/db\_xref="SPTREMBL:Q7XU43"  
/translation="WRLEKAVESWIPAVEKKVDLVSTVDLQVNOQSSAMKT  
EQTPAGTGLPGVPGNGSTGLDGRVQVILNREBAGFTPTIPANGTYASHNVP  
FHLASAMNSFOSGLSVMTGVAVPMTCPEKGDNPIMWKNCEAYDVVGIIHGNVY  
KVATLINFGNAPFLQVRSQLEGVTSSELCEKICTRPNKROQALIRQWIKICOTS  
MEVNDHTTKKEVYVCTMNAFLPMIENYGGCFEDDVLDLQVLDQDEIYVQILQCN  
YGTSSKTHSNPSSSYSHGCELGRRKSPVASVYADLVDBALAREIQMBOULASIS  
IDNNIIEHGKPIASSTSGSNASARPPVYVADIDPNMNYEELQUGAIGTE  
SKGPEVDILPTSTYKIRIFSRKKGDEVCVTCMYKONDRILTKLPCERQVHOTCV  
TKMKIKINVCVCKEYVYSGSK"  
75062..77788  
/gene="OSJNBA0088122.9"  
join(75062..75363,76561..76657,77720..77788)  
/gene="OSJNBA0088122.9"  
/codon\_start=1  
/protein\_id="CAD41577.3"  
/db\_xref="GI:38567913"  
/db\_xref="SPTREMBL:Q7XU44"  
/translation="MSAPSSPHSSSLFSLHLIRTSYARATGNGRESGRROY  
MRRRRSYASPARPPSPGSRWATMAALSLEGAALAAEAERRRVYSGSQRLARGA  
YVDNPPPTKODSNIHSEVNGCSVLCRAMKTNLANCASXPYSEQFPGTPY"  
80941..86909  
/gene="OSJNBA0088122.10"  
join(80941..82090,82295..82536,82597..84075,84346..84484,  
86395..86640,86692..86909)  
/gene="OSJNBA0088122.10"  
/codon\_start=1  
/protein\_id="CAD41578.3"  
/db\_xref="GI:38567914"  
/db\_xref="SPTREMBL:Q7XU43"  
/translation="WRLEKAVESWIPAVEKKVDLVSTVDLQVNOQSSAMKT  
EQTPAGTGLPGVPGNGSTGLDGRVQVILNREBAGFTPTIPANGTYASHNVP  
FHLASAMNSFOSGLSVMTGVAVPMTCPEKGDNPIMWKNCEAYDVVGIIHGNVY  
KVATLINFGNAPFLQVRSQLEGVTSSELCEKICTRPNKROQALIRQWIKICOTS

Query Match  
Best Local Similarity 84.0%; Score 16.8; DB 8; Length 164572;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCATCATGTTGATG 20  
Db 113198 GGTACATACGCATGTTGATG 113217

RESULT 43  
SPNEU908/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

SPNEU908 175936 bp DNA linear HTG 11-JUL-2001  
Streptococcus pneumoniae clone G54, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
AL449930  
AL449930.1 GI:11545155  
HTG; HTGS PHASE2.  
Streptococcus pneumoniae  
Streptococcus pneumoniae  
Bacteria; Filumcutes; Lactobacillales; Streptococcaceae;  
Streptococcus.  
1  
Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Humbert,Y.,  
Friedl,L., Guerrier,M., Grand-Schenk,B., Gandin,C., de  
Francesco,M., Polissi,A., Buell,G., Feger,G., Garcia,B., Peltech,M.  
and Garcia-Buñtos,J.F.  
Annotated draft genomic sequence from a Streptococcus pneumoniae  
type 19F clinical isolate  
Microb. Drug Resist. 7 (2), 99-125 (2001)  
21335329  
11442348  
2 (bases 1 to 175936)  
Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A.,  
Humbert,Y., Friedl,L., Guerrier,M., Grand-Schenk,B., Gandin,C., de  
Francesco,M., Buell,G., Feger,G., Garcia,E., Peltech,M. and  
Garcia-Buñtos,J.F.  
Direct Submission  
Submitted (31-OCT-2000) Research Department, Glaxo Wellcome, S.A.,  
Severo Ochoa 2, 28760 Tres Cantos, SPAIN  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have

\* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 1 175936: contig of 175936 bp in length.  
 Location/Qualifiers  
 1. 175936  
 /organism="Streptococcus pneumoniae"  
 /mol\_type="genomic DNA"  
 /serotype="19F"  
 /db\_xref="taxon:1313"  
 /clone="G54"

ORIGIN

Query Match 84.0%; Score 16.8; DB 2; Length 175936;  
 Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGGCTACTACTGTTGATG 20  
 |||||  
 Db 29955 GGTTCATCTCTGATGATG 29936

RESULT 44  
 AC139894/c  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-278f24, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 70 unordered pieces.  
 AC139894  
 VERSION AC139894.1 GI:28394806  
 KEYWORDS HTG: HTGS PHASE1.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 186184)  
 Muzny,D,Marie, Metzker,M, Lee, Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsprouk, S., Amin, A., Anguiano, D.,  
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Bivaldo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Becorto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guertiz, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,  
 Hollins, B., Howells, S., Huily, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kows, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorenshewa, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
 Mawlin, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,  
 Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,  
 Morris, S., Muidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,  
 Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okunolu, G.,  
 Olajunibegoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H.,  
 Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A.,  
 Popovic, D., Primus, E., Pu, L.-L., Plazo, M., Quiroz, J., Raculin, E.,  
 Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,  
 Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Savery, G., Scherer, S.,  
 Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartbeyn, A.,  
 Sison, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E.,  
 Song, X.-Z., Sorelle, R., Soza, J., Steimle, M., Strong, R., Sutcon, A.,  
 Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,  
 Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D.,  
 Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,  
 Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R.,  
 Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,  
 Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,  
 Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,  
 Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 186184)  
 Worley, K. C.  
 Direct Submission  
 Submitted (16-FEB-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 186184)  
 Worley, K. C.  
 Direct Submission  
 Submitted (27-MAR-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: KOND  
 Center clone name: CH230-278f24  
 ----- Summary Statistics  
 Sequencing Vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 140885 bases at least Q40  
 Consensus quality: 149547 bases at least Q30  
 Consensus quality: 158303 bases at least Q20  
 Estimated insert size: 156705; sum-of-contigs estimation  
 Quality coverage: 2x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a "working draft" sequence. It currently  
 \* consists of 70 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1 1154: contig of 1154 bp in length  
 \* 1  
 1155 1254: gap of unknown length  
 \* 1155 1254: contig of 1029 bp in length  
 \* 1284 2383: gap of unknown length  
 \* 2384 3446: contig of 1063 bp in length  
 \* 3447 3547: gap of unknown length  
 \* 3547 5270: contig of 1723 bp in length  
 \* 5270 5370: gap of unknown length  
 \* 5370 6863: contig of 1493 bp in length  
 \* 6863 6963: gap of unknown length  
 \* 6963 8194: contig of 1232 bp in length  
 \* 8194 8294: gap of unknown length  
 \* 8294 10111: contig of 1817 bp in length  
 \* 10111 10211: gap of unknown length  
 \* 10211 11586: contig of 1375 bp in length  
 \* 11586 11687: gap of unknown length  
 \* 11687 13117: contig of 1431 bp in length  
 \* 13117 13118: gap of unknown length  
 \* 13118 14697: contig of 1480 bp in length



\* 14698 14797: gap of unknown length  
 \* 14798 16013: contig of 1216 bp in length  
 \* 16014 16113: gap of unknown length  
 \* 16114 17514: contig of 1401 bp in length  
 \* 17515 17614: gap of unknown length  
 \* 17615 18912: contig of 1298 bp in length  
 \* 18913 19012: gap of unknown length  
 \* 19013 20041: contig of 1029 bp in length  
 \* 20042 20141: gap of unknown length  
 \* 20142 21580: contig of 1439 bp in length  
 \* 21581 21680: gap of unknown length  
 \* 21681 23564: contig of 1884 bp in length  
 \* 23565 23664: gap of unknown length  
 \* 23665 24769: contig of 1105 bp in length  
 \* 24770 24869: gap of unknown length  
 \* 24870 26701: contig of 1832 bp in length  
 \* 26702 26801: gap of unknown length  
 \* 26802 28355: contig of 1554 bp in length  
 \* 28356 28455: gap of unknown length  
 \* 28456 29847: contig of 1392 bp in length  
 \* 29848 29947: gap of unknown length  
 \* 29948 31194: contig of 1247 bp in length  
 \* 31195 31294: gap of unknown length  
 \* 31295 32553: contig of 1259 bp in length  
 \* 32554 32653: gap of unknown length  
 \* 32654 34999: contig of 2346 bp in length  
 \* 35000 35099: gap of unknown length  
 \* 35100 36223: contig of 1124 bp in length  
 \* 36224 36323: gap of unknown length  
 \* 36324 38111: contig of 1788 bp in length  
 \* 38112 38211: gap of unknown length  
 \* 38212 40164: contig of 1953 bp in length  
 \* 40165 40264: gap of unknown length  
 \* 40265 42432: contig of 2168 bp in length  
 \* 42433 42532: gap of unknown length  
 \* 42533 44066: contig of 1534 bp in length  
 \* 44067 44166: gap of unknown length  
 \* 44167 46233: contig of 2067 bp in length  
 \* 46234 46333: gap of unknown length  
 \* 46334 48587: contig of 2254 bp in length  
 \* 48588 48687: gap of unknown length  
 \* 48688 50591: contig of 1904 bp in length  
 \* 50592 50691: gap of unknown length  
 \* 50692 52837: contig of 2146 bp in length  
 \* 52838 52937: gap of unknown length  
 \* 52939 55123: contig of 2186 bp in length  
 \* 55124 55223: gap of unknown length  
 \* 55224 56867: contig of 1644 bp in length  
 \* 56868 56967: gap of unknown length  
 \* 56968 58897: contig of 1930 bp in length  
 \* 58898 58997: gap of unknown length  
 \* 58999 61198: contig of 2201 bp in length  
 \* 61199 61298: gap of unknown length  
 \* 61299 62820: contig of 1522 bp in length  
 \* 62821 62920: gap of unknown length  
 \* 62921 64789: contig of 1869 bp in length  
 \* 64790 64889: gap of unknown length  
 \* 64890 66789: contig of 1900 bp in length  
 \* 66790 66889: gap of unknown length  
 \* 66890 69075: contig of 2186 bp in length  
 \* 69076 69175: gap of unknown length  
 \* 69176 70324: contig of 1149 bp in length  
 \* 70325 70424: gap of unknown length  
 \* 70425 73643: contig of 3219 bp in length  
 \* 73644 73743: gap of unknown length  
 \* 73744 75986: contig of 2243 bp in length  
 \* 75987 76086: gap of unknown length  
 \* 76087 78980: contig of 2894 bp in length  
 \* 78981 79080: gap of unknown length  
 \* 79081 81809: contig of 2723 bp in length  
 \* 81810 81909: gap of unknown length  
 \* 81910 84998: contig of 3088 bp in length  
 \* 84999 85098: gap of unknown length

\* 85099 88114: contig of 3016 bp in length  
 \* 88115 88214: gap of unknown length  
 \* 88215 91155: contig of 2941 bp in length  
 \* 91156 91255: gap of unknown length  
 \* 91256 93241: contig of 1986 bp in length  
 \* 93242 93341: gap of unknown length

Query Match 84.0%; Score 16.8; DB 2; length 186184;  
 Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20  
 Db 169681 GGCCTACTCATGTTGATG 169662

RESULT 45  
 AC119611/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE  
 AUTHORS  
 1 (bases 1 to 209706)  
 Muzny,D,Marie, Metczer,M, Lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Albrooks,S, Amun,A, Anguiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Dexamio,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howells,S, Huix,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpachy,L, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,Z, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshewa,L, Louised,H, Lorado,R,S, Lu,X, Ma,J, Maheshwari,M, Mahimuddin,M, Mahmoud,M, Mallory,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawlinyevic,A, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwankweliemh,O, Okwomu,G, Olarinmagaon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,B, Pu,L, L, Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Reiter,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojase,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shartsbeyn,A, Sisson,I, Sitter,C,D, Snares,D, Sneed,A, Sodergren,E, Song,X-Z, Sorille,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,



Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K.,  
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
Weinstock,G. and Gibbs,R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 209706)  
Worley,K.C.  
Direct Submission  
Submitted (29-APR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 209706)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23808767.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu

----- Project Information  
Center project name: GYON  
Center clone name: CH230-440C15

----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 187370 bases at least Q40  
Consensus quality: 188624 bases at least Q30  
Consensus quality: 189405 bases at least Q20  
Estimated insert size: 192639; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* (Note: This is a 'working draft' sequence. It currently  
\* consists of 1 contig. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 209706: contig of 209706 bp in length.  
Location/Qualifiers  
1. 209706  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-440C15"  
2286. 3117  
/note="clone boundary  
clone\_end:5p6  
site:  
end\_sequence:B2128438"  
202827. 203690  
/note="clone boundary

FEATURES  
source  
misc\_feature  
misc\_feature  
misc\_feature

clone\_end:77  
site:  
end\_sequence:B2128438"  
206213. 207242  
/note="wgs\_end\_extension  
clone\_end:77"  
207915. 209706  
/note="wgs\_end\_extension  
clone\_end:77"

ORIGIN  
Query Match 84.0%; Score 16.8; DB 2; Length 209706;  
Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
Db 98490 GGTCCATCTCATGTTGATG 98471

RESULT 46  
AC129598/c 213381 bp DNA linear ROD 25-NOV-2003  
LOCUS  
DEFINITION  
Mus musculus BAC clone RP23-28305 from chromosome 18, complete  
sequence.  
AC129598  
AC129598.4 GI:33285248  
HTG.  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 213381)  
Isak,A. and Haglund,K.  
The sequence of Mus musculus BAC clone RP23-28305  
Unpublished (2001)  
2 (bases 1 to 213381)  
Wilson,R.  
Sequencing of Mus musculus  
Unpublished (2001)  
3 (bases 1 to 213381)  
McPherson,J.D. and Waterston,R.H.  
Direct Submission  
Submitted (30-JUL-2002) Genome Sequencing Center, 4444 Forrest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 213381)  
Wilson,R.K.  
Direct Submission  
Submitted (20-MAY-2003) Genome Sequencing Center, 4444 Forrest Park  
Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 213381)  
Wilson,R.K.  
Direct Submission  
Submitted (26-JUL-2003) Genome Sequencing Center, 4444 Forrest Park  
Parkway, St. Louis, MO 63108, USA  
6 (bases 1 to 213381)  
Wilson,R.  
Direct Submission  
Submitted (25-NOV-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Jul 26, 2003 this sequence version replaced gi:30911163.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: submissions@wuston.wustl.edu

----- Summary Statistics  
Center project name: M\_BA0283005

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap

between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
<http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RPCI-23 BAC Library has been constructed by Kazutoyo Oseigawa  
and Minako Tateno in the laboratory of Pieter de Jong  
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or  
brain genomic DNA. The clone and detailed information can be  
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or  
Pieter de Jong and coworkers at <http://www.chori.org>

#### NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

#### FEATURES

##### source

```
1..213381
  location/Qualifiers
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="18"
    /map="18"
    /clone="RPCI-28305"
    /clone_1lb="RPCI-23"
    61..124
    /rpt_family="MER1_type"
    236..338
    /rpt_family="L1"
    3234..3467
    /rpt_family="L1"
    4674..4747
    /rpt_family="Alu"
    5793..5930
    /rpt_family="MER1_type"
    6211..6610
    /rpt_family="L1"
    6804..7142
    /rpt_family="L1"
    9876..10022
    /rpt_family="Alu"
    11296..11350
    /rpt_family="ERV1"
    11818..12016
    /rpt_family="L1"
    12019..13959
    /rpt_family="L1"
    13961..14028
    /rpt_family="L1"
    14028..14346
    /rpt_family="L1"
    14362..14699
    /rpt_family="MALR"
    14704..14809
    /rpt_family="L1"
    14832..15990
    /rpt_family="L1"
    15989..16163
    /rpt_family="L1"
    16170..16379
    /rpt_family="L1"
    16380..16811
    /rpt_family="MALR"
```

```
repeat_region 16831..17134
               /rpt_family="L1"
repeat_region 17112..17202
               /rpt_family="L1"
repeat_region 17195..17789
               /rpt_family="L1"
unsure 19446..19613
       /note="Sequence derived from one plasmid subclone."
repeat_region 24760..24858
               /rpt_family="L1"
repeat_region 24882..26242
               /rpt_family="L1"
repeat_region 26273..27217
               /rpt_family="L1"
repeat_region 27233..27300
               /rpt_family="B2"
repeat_region 27863..27911
               /rpt_family="B2"
repeat_region 28816..29068
               /rpt_family="L2"
repeat_region 31173..31873
               /rpt_family="RMR13B"
repeat_region 32908..33503
               /rpt_family="L1"
repeat_region 38155..39637
               /rpt_family="L1"
repeat_region 42874..42931
               /rpt_family="L1"
repeat_region 42985..43199
               /rpt_family="MER1_type"
repeat_region 43419..43722
               /rpt_family="RMR6B"
repeat_region 43742..44280
               /rpt_family="RMR6A"
repeat_region 45706..45906
               /rpt_family="RMR6A"
repeat_region 46062..46394
               /rpt_family="RMR17C"
repeat_region 46258..46627
               /rpt_family="ERVK"
misc_feature 48421..48648
             /note="Sequence derived from PCR product of project DNA."
repeat_region 48480..48877
               /rpt_family="RMR17C"
repeat_region 49189..49637
               /rpt_family="ERVK"
repeat_region 49955..50433
               /rpt_family="MALR"
repeat_region 50852..50962
               /rpt_family="L1"
repeat_region 50965..51177
               /rpt_family="L1"
repeat_region 51200..52281
               /rpt_family="L1"
repeat_region 52304..52442
               /rpt_family="Alu"
repeat_region 52838..52889
               /rpt_family="Alu"
repeat_region 52934..53198
               /rpt_family="L1"
repeat_region 53908..54059
               /rpt_family="B4"
repeat_region 56035..56662
               /rpt_family="L1"
unsure 56230..56337
       /note="Sequence derived from one plasmid subclone."
repeat_region 56660..57273
               /rpt_family="L1"
repeat_region 57429..57577
               /rpt_family="Alu"
repeat_region 57628..58265
               /rpt_family="L1"
repeat_region 59055..59425
```

```

repeat_region      /rpt_family="L1"
                   59626..60201
                   /rpt_family="L1"

Query Match      84.0%; Score 16.8; DB 10; Length 213381;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GGTCCACTCATGTTGATG 20
    |||||
Db 104478 GGTCAACACTCATGTTGATG 104459

RESULT 47
AC107756 216911 bp DNA linear ROD 30-SEP-2003
LOCUS Mus musculus chromosome 18, clone RP23-214K19, complete sequence.
AC107756
AC107756.9 GI:33414530
VERSION HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 216911)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
JOURNAL Mus musculus chromosome 18, clone RP23-214K19
TITLE Unpublished
AUTHORS 2 (bases 1 to 216911)
          Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
          Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckhalter,B.,
          Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
          Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
          Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,
          Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
          Gao,S., Goid,S., Goyette,M., Graham,L., Grand-Pierre,N.,
          Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
          Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
          Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
          Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
          McEwan,P., McKernan,K., Meldrim,J., Menus,L., Mihova,T.,
          Mlegana,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
          Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
          Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
          Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
          Roestli,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
          Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
          Straus,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,
          Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
          Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
          Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 216911)
          Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
          Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
          Boguslavsky,L., Bouckhalter,B., Camarata,J., Chang,J., Choepel,Y.,
          Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
          Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
          Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
          Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
          Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
          Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
          Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
          Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
          Meldrim,J., Menus,L., Mihova,T., Mlegana,V., Murphy,T., Naylor,J.,
          Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
          O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
          Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,
          Smith,C.,
          Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
          Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M.,
          Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
          Wyman,D., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.

```

```

TITLE Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
JOURNAL Wyman,D., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.
REFERENCE Direct Submission
AUTHORS Submitted (25-JUN-2003) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
          4 (bases 1 to 216911)
          Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
          Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
          Boguslavsky,L., Bouckhalter,B., Camarata,J., Chang,J., Choepel,Y.,
          Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
          Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
          Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
          Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
          Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
          Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
          Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
          Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
          Meldrim,J., Menus,L., Mihova,T., Mlegana,V., Murphy,T., Naylor,J.,
          Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
          O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
          Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,
          Smith,C.,
          Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
          Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M.,
          Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
          Wyman,D., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2003) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 5 (bases 1 to 216911)
          Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
          Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
          Boguslavsky,L., Bouckhalter,B., Camarata,J., Chang,J., Choepel,Y.,
          Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
          Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fero,S.,
          Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
          Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
          Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
          Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
          Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C.,
          Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
          Meldrim,J., Menus,L., Mihova,T., Mlegana,V., Murphy,T., Naylor,J.,
          Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
          O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
          Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P.,
          Smith,C.,
          Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
          Talamas,J., Testfaye,S., Theodore,J., Topham,K., Travers,M.,
          Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
          Wyman,D., Young,G., Zainoun,J., Zemdek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2003) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS On Aug 3, 2003 this sequence version replaced gi:32189542.
          All repeats were identified using RepeatMasker:
          Smit, A.P.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          -----
          Center: Whitehead Institute/ MIT Center for Genome Research
          Center code: WtBR
          Web site: http://www-seq.wi.mit.edu
          Contact: sequence_submissions@genome.wi.mit.edu
          -----
          Project Information
          Center project name: L20399
          Center clone name: 214_K_19
          -----
          Location/Qualifiers
          1..216911
          /organism="Mus musculus"
          /mol_type="genomic DNA"
          /db_xref="taxon:10090"
          /chromosome="18"
          /map="18"

```

## FEATURES

source

misc_feature	/clone="RP23-214K1.9" 1. .2633	/clone_1lb="RPCT-23 Female Mouse BAC"
repeat_region	/note="wgs_end_extension clone_end:SP6" 1. .783	
repeat_region	/rpt_family="L1a" 800. .2024	
repeat_region	/rpt_family="L1" 2034. .3692	
misc_feature	/rpt_family="L1_MM" 2634. .2639	
	/note="clone boundary clone_end:SP6 site:ECOR1"	
repeat_region	complement(3908. .4155) /rpt_family="L1"	
repeat_region	4156. .4196 /rpt_family="TR)n"	
repeat_region	complement(4197. .4639) /rpt_family="L1a"	
repeat_region	complement(4657. .5355) /rpt_family="LX5"	
repeat_region	complement(5372. .5778) /rpt_family="L1"	
repeat_region	5826. .5890 /rpt_family="TRATATG)n"	
repeat_region	complement(6783. .6901) /rpt_family="MTR3"	
repeat_region	complement(7147. .7655) /rpt_family="L1_MM"	
repeat_region	complement(7852. .7962) /rpt_family="PB1D10"	
unsure	8947. .9105 /note="single clone coverage"	
repeat_region	complement(9226. .9405) /rpt_family="MIR"	
repeat_region	complement(9460. .9819) /rpt_family="RLRERN_MM"	
repeat_region	complement(9820. .9976) /rpt_family="MERSA"	
repeat_region	12506. .12585 /rpt_family="TC)n"	
repeat_region	complement(13122. .13294) /rpt_family="LX2"	
repeat_region	13295. .14013 /rpt_family="RLR13C"	
repeat_region	14017. .14180 /rpt_family="RLR13D"	
repeat_region	complement(14181. .14269) /rpt_family="LX"	
repeat_region	14618. .14756 /rpt_family="B3"	
repeat_region	complement(14856. .15358) /rpt_family="L1"	
repeat_region	16436. .16848 /rpt_family="RLR10C"	
repeat_region	16949. .17012 /rpt_family="TRTC)n"	
repeat_region	complement(17015. .17261) /rpt_family="L1_MM"	
repeat_region	17480. .17686 /rpt_family="LX8"	
repeat_region	17921. .19038 /rpt_family="LX7"	
repeat_region	complement(19022. .19450) /rpt_family="L1"	
repeat_region	19451. .20633 /rpt_family="L1_MM"	

Query Match	84.0%;	Score 16.8;	DB 10;	Length 216911;
Best Local Similarity	90.0%;	Pred. No. 1.1e+03;		
Matches 18;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

```

QY      1  GGTCCATCTCATGTTGATG  20
          |||||  |||||  |||||
Db      195618  GGTCAACACTCATGTTGATG  195637

```

RESULT	48		
AC107742			
LOCUS			
DEFINITION	Mus musculus clone RP23-331D7, WORKING DRAFT SEQUENCE, 9 ordered		
	AC107742	225182 bp	DNA linear HTG 07-JUN-2003

VERSION	AC1077442.3	GI:31455675
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS	(pages 1 to 225182) Birren, B., Nusbaum, C. and Lander, E. Mus musculus, clone RP23-331D7 Unpublished 2 (pages 1 to 225182) Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., P.

Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gager, D., Galagan, J., Gaidyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazars, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Martinus, N., Matthews, C., McCarthy, M., McKernan, P., McKernan, K., Meldrum, J., Mensu, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., North, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pletier, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topfman, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zahoun, J., Zemek, L., Zimmer, A. and Zody, M.

REFERENCE  
3 (bases 1 to 225182)  
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

Collamore, A., Cook, A., Cooke, P., Corum, B., Deavellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, T., Faro, S., Ferreira, P., Filizcerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Haez, N., Haggopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, T., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Poh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, T., Matthews, C., McCarthy, M., Meltrin, T., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, C., Plunkhang, P., Piere, N., Rachtupara, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassilev, H., Venkataraman, V.S., Viet, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE  
JOURNAL  
Submitted (07-JUN-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT  
On Jun 7, 2003 this sequence version replaced gi|20303769.

All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

Project Information

Center project name: L20349

Center clone name: 331\_D7

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 223109 bases at least Q40

Consensus quality: 223996 bases at least Q30

Consensus quality: 224258 bases at least Q20

Insert size: 206000; agarose-fp

Insert size: 224382; sum-of-contigs

Quality coverage: 8.6 in Q20 bases; agarose-fp

Quality coverage: 7.9 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\*  
\* 1 23186: contig of 23186 bp in length  
\* 23187 23286: gap of 100 bp  
\* 23287 27308: contig of 4022 bp in length  
\* 27309 27408: gap of 100 bp  
\* 27409 36046: contig of 8638 bp in length  
\* 36047 36147: gap of 100 bp  
\* 36147 50755: contig of 14609 bp in length  
\* 50756 50855: gap of 100 bp  
\* 50856 69982: contig of 19127 bp in length  
\* 69983 70083: gap of 100 bp  
\* 70083 146450: contig of 76368 bp in length  
\* 146451 146550: gap of 100 bp  
\* 146551 175621: contig of 29071 bp in length  
\* 175622 175721: gap of 100 bp  
\* 175722 207711: contig of 31990 bp in length  
\* 207712 207811: gap of 100 bp  
\* 207812 225182: contig of 17371 bp in length.

#### FEATURES

source

1..225182  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-331D7"  
/clone\_lib="RP23-331D7"  
1..23186  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
23287..27308  
/note="assembly\_fragment"  
27409..36046  
/note="assembly\_fragment"  
36147..50755  
/note="assembly\_fragment"  
50856..69982  
/note="assembly\_fragment"  
70083..146450  
/note="assembly\_fragment"  
146551..175621  
/note="assembly\_fragment"  
175722..207711  
/note="assembly\_fragment"

#### ORIGIN

Query Match 84.0% Score 16.8; DB 2; Length 225182;  
Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTTGATG 20  
|||||  
Db 38164 GGTCCAACTCAGTTGATG 38183

RESULT 49  
AE016936/c 300340 bp DNA linear BCT 28-MAR-2003

LOCUS

DEFINITION Bacteroides thetaiotaomicron VPI-5482, section 11 of 21 of the

complete genome.

ACCESSION AE016936 AE015928

VERSION AE016936.1 GI:29339725

KEYWORDS

SOURCE

ORGANISM

Bacteroides thetaiotaomicron VPI-5482

Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;

Bacteroidaceae; Bacteroides.

REFERENCE Xu, J., Bjurvell, M.K., Himrod, J., Deng, S., Carmichael, L.K.,

Chiang, H.C., Hooper, L.V. and Gordon, J.I.

A Genomic View of the Human-Bacteroides thetaiotaomicron Symbiosis

Science 299 (5615), 2074-2076 (2003)

JOURNAL 12663928

PUBMED 2 (bases 1 to 300340)

AUTHORS Xu, J., Bjurvell, M.K., Himrod, J., Deng, S., Carmichael, L.K.,

Chiang, H.C., Hooper, L.V. and Gordon, J.I.

Direct Submission

Submitted (31-OCT-2002) Department of Molecular Biology and

Pharmacology, Washington University in St. Louis, 660 S. Euclid,

St. Louis, MO 63110, USA

Location/Qualifiers

1..300340

/organism="Bacteroides thetaiotaomicron VPI-5482"

/mol\_type="genomic DNA"

/strain="VPI-5482"

/db\_xref="taxon:226186"

91..321

/locus\_tag="BT2414"

91..321

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"

/locus\_tag="BT2414"



Query Match 84.0%; Score 16.8; DB 1; Length 300340;  
 Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTGATG 20  
 |||||  
 DB 114904 GGTCATGCTCATGTGATG 114885

## RESULT 50

AX571760/c AX571760 349980 bp DNA linear PAT 29-NOV-2002  
 LOCUS  
 DEFINITION Sequence 4979 from Patent WO02077021.  
 ACCESSION AX571760  
 VERSION AX571760.1 GI:26003952  
 KEYWORDS

SOURCE Streptococcus pneumoniae  
 ORGANISM Streptococcus pneumoniae  
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 Streptococcus.

## REFERENCE

1 Masignani,V., Tettelin,H. and Fraser,C.  
 Streptococcus pneumoniae proteins and nucleic acids  
 TITLE Patent: WO 02077021-A 4979 03-OCT-2002;  
 JOURNAL Chiron Spa (IT) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)  
 Location/Qualifiers

## FEATURES

source 1. 349980  
 /organism="Streptococcus pneumoniae"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:1313"  
 /note="seq 4979 too long: 2.162.598 bases-replaced by  
 following seq:-seq 4979: from 0.000.001 to 0.349.980-seq  
 4980: from 0.300.001 to 0.649.980-seq 4981: from 0.600.001  
 to 0.949.980-seq 4982: from 0.900.001 to 1.249.980-seq  
 4983: from 1.200.001 to 1.549.980-seq 4984: from 1.500.001  
 to 1.849.980-seq 4985: from 1.800.001 to 2.149.980-seq  
 4986: from 2.100.001 to 2.162.598"

## ORIGIN

Query Match 84.0%; Score 16.8; DB 6; Length 349980;  
 Best Local Similarity 90.0%; Pred. No. 1.1e+03;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTGATG 20  
 |||||  
 DB 67573 GGTCATGCTCATGTGATG 67554

Search completed: August 17, 2004, 15:02:53  
 Job time : 1612 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 13:18:52 ; Search time 360 Seconds  
(without alignments)  
236.011 Million cell updates/sec

Title: US-09-825-489-3  
Perfect score: 20  
Sequence: 1 ggtccatcactcatgtgatg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 800 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: geneeqn1980s:\*
- 2: geneeqn1990s:\*
- 3: geneeqn2000s:\*
- 4: geneeqn2001as:\*
- 5: geneeqn2001bs:\*
- 6: geneeqn2002s:\*
- 7: geneeqn2003as:\*
- 8: geneeqn2003bs:\*
- 9: geneeqn2003cs:\*
- 10: geneeqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	4	AAD21384 Antisense
2	20	100.0	486	6	ABL84566 Human ova
3	20	100.0	506	6	ABL79139 Human ova
4	20	100.0	513	6	ABL79129 Human ova
5	20	100.0	822	4	AAS63227 Human xer
6	20	100.0	1377	4	AAD21394 Human xer
7	20	100.0	1407	4	AAS34872 CDNA enco
8	20	100.0	1407	6	ADC46030 Human neo
9	20	100.0	4670	6	ABZ35476 Human gen
10	16.8	84.0	1029	7	ABX05799 S. pneumo
11	16.8	84.0	14273	2	AAVS2173 Streptoco
12	16.8	84.0	110000	7	ABSS56454_00 Streptoco
13	16.4	82.0	6765	7	ABT13966 Human Ctr
14	16.4	82.0	15246	7	ABT13954 Human Ctr
15	16.4	82.0	110000	6	ABA90521_15 Continuation (16 o
16	15.8	79.0	431	6	ABV97710 Human pan
17	15.8	79.0	462	3	ACH25004 Human adu
18	15.8	79.0	614	3	AAA16057 Human col
19	15.8	79.0	2830	4	ABU23928 Drosophil
20	15.8	79.0	6454	4	ABU11966 Drosophil
21	15.8	79.0	8114	4	ABU03787 Drosophil
22	15.8	79.0	18737	4	ABU03786 Drosophil
23	15.8	79.0	24401	4	ABU03396 Drosophil

24	15.4	77.0	551	4	ABK42099 CDNA enco
25	15.4	77.0	551	8	ADB59766 Connectiv
26	15.4	77.0	609	5	ABV55825 Human pro
27	15.4	77.0	633	4	AAH52789 S. epider
28	15.4	77.0	675	6	AAH53645 S. epider
29	15.4	77.0	675	6	ABN90925 Staphyloc
30	15.4	77.0	2132	3	AAC46198 Arabidops
31	15.4	77.0	2135	3	AAC40444 Arabidops
32	15.4	77.0	2946	4	AAH54374 S. epider
33	15.4	77.0	3021	4	AAH54380 S. epider
34	15.4	77.0	3391	4	AAH54393 S. epider
35	15.4	77.0	7722	5	AAH85546 DNA encod
36	15.4	77.0	28564	9	ADD46613 Human gen
37	15.4	77.0	28564	9	ADD46613 Human gen
38	15.4	77.0	335913	5	AAI61371 Soybean 2
39	15.4	77.0	335913	5	AAI61372 Soybean 2
40	15.2	76.0	308	2	AAV75448 Staphyloc
41	15.2	76.0	330	2	AAV75448 Staphyloc
42	15.2	76.0	443	7	AAD50058 Human fat
43	15.2	76.0	468	3	AAFI1349 Aspergill
44	15.2	76.0	472	7	AAD50056 Human fat
45	15.2	76.0	499	6	ABK69101 DNA encod
46	15.2	76.0	499	7	AAD50057 Human fat
47	15.2	76.0	547	6	ABK69131 DNA encod
48	15.2	76.0	548	7	AAD50054 Human fat
49	15.2	76.0	578	3	AAD50053 Human fat
50	15.2	76.0	685	7	AAFI1457 Aspergill
51	15.2	76.0	708	7	AAD50052 Human fat
52	15.2	76.0	785	4	AAK72968 Human imm
53	15.2	76.0	912	7	AAZ51334 Aspergill
54	15.2	76.0	919	6	AAFI1292 Chlamydia
55	15.2	76.0	1035	6	ABU191292 Chlamydia
56	15.2	76.0	1107	7	ACA38617 Prokaryot
57	15.2	76.0	1287	5	AAH67093 C. glutami
58	15.2	76.0	1335	4	AAFI68053 Corynebact
59	15.2	76.0	1410	6	AAFI68052 Corynebact
60	15.2	76.0	2000	6	ABZ17137 Arabidops
61	15.2	76.0	2000	7	ADA68826 Arabidops
62	15.2	76.0	2368	3	AAZ49575 S. aureus
63	15.2	76.0	2385	2	AAQ27956 Cholester
64	15.2	76.0	2608	3	AAZ49578 S. aureus
65	15.2	76.0	2916	8	ADA02611 Mouse NfK
66	15.2	76.0	2916	9	ADB72349 Mouse NfK
67	15.2	76.0	2916	9	ADB82943 Mouse NfK
68	15.2	76.0	3032	5	ABV25508 Human pro
69	15.2	76.0	3324	7	AAD50048 Human fat
70	15.2	76.0	3334	6	ABK35711 CDNA sequ
71	15.2	76.0	3657	7	AAD50047 Human fat
72	15.2	76.0	3663	7	AAD50046 Human fat
73	15.2	76.0	3670	9	ADB75301 Prostate
74	15.2	76.0	3762	7	AAD50064 Human fat
75	15.2	76.0	3768	7	AAD50063 Human fat
76	15.2	76.0	3892	6	ABU199776 Mouse ISG
77	15.2	76.0	3892	6	ADA02610 Mouse NfK
78	15.2	76.0	3892	9	ADB72348 Mouse NfK
79	15.2	76.0	3892	9	ADB82942 Mouse NfK
80	15.2	76.0	4884	8	ADA31191 DNA encod
81	15.2	76.0	5327	4	AAFI30428 Vesicular
82	15.2	76.0	5674	4	ABU126242 Drosophil
83	15.2	76.0	5881	9	ADD27432 Human adi
84	15.2	76.0	5881	9	ADD27432 Human adi
85	15.2	76.0	6222	8	ADA02758 Human MOR
86	15.2	76.0	6222	9	ADB72496 Human MOR
87	15.2	76.0	6222	9	ADB85238 Human MOR
88	15.2	76.0	6330	8	AAFI30429 Vesicular
89	15.2	76.0	6537	8	ADA02757 Human MOR
90	15.2	76.0	6537	9	ADB72495 Human MOR
91	15.2	76.0	6537	9	ADB85237 Human MOR
92	15.2	76.0	7798	9	ADB31351 Human dia
93	15.2	76.0	8354	4	ABU02619 Drosophil
94	15.2	76.0	14311	2	AAFI38744 Plasmid p
95	15.2	76.0	14311	5	AAFI38744 Plasmid p
96	15.2	76.0	14311	5	AAFI38744 Plasmid p

97	15.2	76.0	14361	4	ABE02618	AbI02618 Drosophill	c 170	14.8	74.0	1165	6	ABS54004	AbS54004 cDNA enco
98	15.2	76.0	17173	7	ABZ80234	Abz80234 Mouse tra	c 171	14.8	74.0	1201	7	ABT33402	ABT33402 NOVA DNA
99	15.2	76.0	36412	9	ABD74383	Abd74383 Mycobacte	c 172	14.8	74.0	1242	7	AAF90254	AAf90254 DNA encod
c 100	15.2	76.0	82993	6	ABX09140	Abx09140 Mycobacte	173	14.8	74.0	1242	5	AAST7861	AAst7861 DNA encod
c 101	15.2	76.0	96598	8	ADN02609	Adn02609 Mouse NfK	174	14.8	74.0	1242	6	AAD21891	AA221891 Human tum
c 102	15.2	76.0	96598	8	ADB72347	Adb72347 Mouse NfK	c 175	14.8	74.0	1357	6	AUC86072	AUC86072 Human GPC
c 103	15.2	76.0	96598	9	ADBE2941	Adbe2941 Mouse NfK	c 176	14.8	74.0	1357	6	ADC86070	ADC86070 Human GPC
c 104	15.2	76.0	110000	2	AAK91990_01	ContInnation (2 of	c 177	14.8	74.0	1364	6	ADC86070	ADC86070 Human GPC
c 105	15.2	76.0	110000	4	AAI99682_21	ContInnation (22 o	c 178	14.8	74.0	1568	3	AAAC98032	AAc98032 Human col
c 106	15.2	76.0	110000	4	AAI99682_21	ContInnation (22 o	c 179	14.8	74.0	1719	7	ACA52604	ACa52604 Prokaryot
c 107	15.2	76.0	326014	6	ABK89286	Abk89286 Human gen	c 180	14.8	74.0	1818	7	ACA31066	ACa31066 Prokaryot
c 108	15.2	76.0	349980	5	AAH68530	Aah68530 c glutamI	c 181	14.8	74.0	1836	4	AAI19464	AAI19464 Probe #93
c 109	15	75.0	593	4	AAI17156	AAI17156 Probe #70	c 182	14.8	74.0	1836	4	ABA64480	ABa64480 Human foe
c 110	15	75.0	593	4	ABA61657	AbA61657 Human foe	c 183	14.8	74.0	1836	4	AAI44657	AAI44657 Probe #13
c 111	15	75.0	593	4	AAI41569	AAI41569 Probe #10	c 184	14.8	74.0	1836	4	ABA31619	ABa31619 Probe #10
c 112	15	75.0	593	4	ABA29315	AbA29315 Probe #77	c 185	14.8	74.0	1836	4	AAK12937	AAK12937 Human bra
c 113	15	75.0	593	4	AAK35852	AaK35852 Human bon	c 186	14.8	74.0	1836	4	ABS38231	ABs38231 Human liv
c 114	15	75.0	593	4	AAK09958	AaK09958 Human bra	c 187	14.8	74.0	1836	6	ABS12734	ABs12734 Human gen
c 115	15	75.0	593	4	ABS35562	AbS35562 Human liv	c 188	14.8	74.0	1859	7	ABZ57038	ABz57038 Human mII
c 116	15	75.0	593	6	ABS10085	AbS10085 Human gen	c 189	14.8	74.0	1945	4	ABL16181	ABl16181 Drosophill
c 117	15	75.0	761	9	ADD17343	Add17343 DNA (SegI	c 190	14.8	74.0	1987	4	AAI13905	AAI13905 Probe #38
c 118	15	75.0	1851	9	ADD43533	Add43533 Mouse ope	c 191	14.8	74.0	1987	4	ABA55617	ABa55617 Human foe
c 119	15	75.0	2000	9	ADD43525	Add43525 Mouse cDN	c 192	14.8	74.0	1987	4	AAI35272	AAI35272 Probe #39
c 120	14.8	74.0	30	6	ABX69902	Abx69902 Novel HeI	c 193	14.8	74.0	1987	4	ABA45139	ABa45139 Human bre
c 121	14.8	74.0	100	7	ACD78987	AcD78987 E. coli K	c 194	14.8	74.0	1987	4	ABA25315	ABa25315 Probe #37
c 122	14.8	74.0	138	7	ACA14155	ACa14155 Prokaryot	c 195	14.8	74.0	1987	4	AAK29312	AAK29312 Human bon
c 123	14.8	74.0	276	4	AAK63634	AaK63634 Human imm	c 196	14.8	74.0	1987	4	AAK03845	AAK03845 Human bra
c 124	14.8	74.0	315	8	ADA29841	AdA29841 DNA encod	c 197	14.8	74.0	1987	5	ABS28932	ABs28932 Human liv
c 125	14.8	74.0	348	8	ADA29940	AdA29940 DNA encod	c 198	14.8	74.0	1987	5	AAI03774	AAI03774 Probe #37
c 126	14.8	74.0	380	5	AAI05027	AAI05027 Probe #50	c 199	14.8	74.0	1987	6	ABS03870	ABs03870 Human gen
c 127	14.8	74.0	380	5	AAI09998	AAI09998 Probe #99	c 200	14.8	74.0	2000	7	ADA72679	ADa72679 Rice gene
c 128	14.8	74.0	380	5	ABS05391	AbS05391 Human gen	c 201	14.8	74.0	2033	6	ABS51296	ABs51296 Human REM
c 129	14.8	74.0	380	5	ABK17888	Abk17888 Human gen	c 202	14.8	74.0	2379	6	ABZ14366	ABz14366 Arabidops
c 130	14.8	74.0	414	7	ACA19003	AcA19003 Prokaryot	c 203	14.8	74.0	2737	2	AAK90784	AAK90784 Mucant DN
c 131	14.8	74.0	459	4	AAI10174	AAI10174 Probe #10	c 204	14.8	74.0	2737	2	AAK90780	AAK90780 PCR ampli
c 132	14.8	74.0	459	4	ABAS1807	AbA51807 Human foe	c 205	14.8	74.0	2737	2	AAK90785	AAK90785 Detergente
c 133	14.8	74.0	459	4	AAI31424	AAI31424 Probe #11	c 206	14.8	74.0	2814	4	AAH57377	AAH57377 Human hea
c 134	14.8	74.0	459	4	ABA21636	AbA21636 Probe #14	c 207	14.8	74.0	3115	4	ABL18346	ABl18346 Drosophill
c 135	14.8	74.0	459	4	AAK00114	AaK00114 Human bra	c 208	14.8	74.0	3153	6	ABN87601	ABn87601 Human cop
c 136	14.8	74.0	459	4	ABS25121	AbS25121 Human liv	c 209	14.8	74.0	3295	9	ADB68998	ADb68998 C. neofor
c 137	14.8	74.0	459	6	ABS00120	AbS00120 Human gen	c 210	14.8	74.0	3869	4	AAK79858	AAK79858 Human imm
c 138	14.8	74.0	540	5	ABS72578	AbS72578 DNA encod	c 211	14.8	74.0	4042	9	ADE07449	ADe07449 Novel cod
c 139	14.8	74.0	546	8	ADA29705	AdA29705 DNA encod	c 212	14.8	74.0	4125	7	ADA70202	ADa70202 Rice gene
c 140	14.8	74.0	577	4	AAI23122	AAI23122 Probe #13	c 213	14.8	74.0	4212	7	ABL16180	ABl16180 Drosophill
c 141	14.8	74.0	577	4	ABA68215	AbA68215 Human foe	c 214	14.8	74.0	4595	6	ABK63624	ABK63624 Rat seque
c 142	14.8	74.0	577	4	AAI48430	AAI48430 Probe #17	c 215	14.8	74.0	4595	9	ADB58031	ADB58031 Toxicity-
c 143	14.8	74.0	577	4	ABA50275	AbA50275 Human bre	c 216	14.8	74.0	4595	9	ADB52503	ADB52503 Primary r
c 144	14.8	74.0	577	4	ABA35227	AbA35227 Probe #13	c 217	14.8	74.0	5438	5	ABV24305	ABv24305 Human pro
c 145	14.8	74.0	577	4	AAK42351	AaK42351 Human bon	c 218	14.8	74.0	7047	6	AAI49486	AAI49486 Drosophill
c 146	14.8	74.0	577	4	AAK16596	AaK16596 Human bra	c 219	14.8	74.0	7047	6	ABN85738	ABn85738 Drosophill
c 147	14.8	74.0	577	4	ABS41960	AbS41960 Human liv	c 220	14.8	74.0	7156	2	AAV15586	AAV15586 Lactobaci
c 148	14.8	74.0	577	5	AAI08777	AAI08777 Probe #87	c 221	14.8	74.0	7448	4	ABL05347	ABl05347 Drosophill
c 149	14.8	74.0	577	6	ABS16406	AbS16406 Human gen	c 222	14.8	74.0	7566	4	AAI07322	AAI07322 Human rep
c 150	14.8	74.0	635	3	AAFO8438	AaFO8438 Fusarium	c 223	14.8	74.0	7601	4	AAK81252	AAK81252 Human imm
c 151	14.8	74.0	642	7	ACA31472	ACa31472 Prokaryot	c 224	14.8	74.0	7741	7	ABX97045	ABx97045 Human NOV
c 152	14.8	74.0	784	4	AAK85396	AaK85396 Human imm	c 225	14.8	74.0	8008	8	AAAD51565	AAAd51565 Human str
c 153	14.8	74.0	858	6	ABZ43136	AbZ43136 Human GPC	c 226	14.8	74.0	11656	4	ABL05346	ABl05346 Drosophill
c 154	14.8	74.0	858	6	ABZ43134	AbZ43134 Human GPC	c 227	14.8	74.0	13954	2	AAK20641	AAK20641 Polynucle
c 155	14.8	74.0	944	6	ABU92761	AbI92761 G proteIn	c 228	14.8	74.0	14668	4	ABL28811	ABl28811 Drosophill
c 156	14.8	74.0	954	4	AAH31607	AaH31607 Human oIf	c 229	14.8	74.0	15246	7	ABT13955	ABt13955 Simian VI
c 157	14.8	74.0	954	4	AAH32129	AAH32129 Human oIf	c 230	14.8	74.0	18060	4	ABL14126	ABl14126 Drosophill
c 158	14.8	74.0	957	5	AAAS42395	AAaS42395 Human cDN	c 231	14.8	74.0	19174	4	AAK81253	AAK81253 Human imm
c 159	14.8	74.0	957	5	AAAS42396	AAaS42396 Human cDN	c 232	14.8	74.0	29283	9	ADC87182	ADc87182 Human GPC
c 160	14.8	74.0	957	6	ABK68559	AbK68559 Human DNA	c 233	14.8	74.0	29781	4	ABL28810	ABl28810 Drosophill
c 161	14.8	74.0	957	6	ABK68557	AbK68557 Human DNA	c 234	14.8	74.0	31203	4	AAK81254	AAK81254 Human imm
c 162	14.8	74.0	957	6	ABK37682	AbK37682 DNA encod	c 235	14.8	74.0	32167	5	ABA20621	ABa20621 Human ner
c 163	14.8	74.0	972	6	ABK37681	AbK37681 DNA encod	c 236	14.8	74.0	33675	6	AAD31202	AA31202 Oryza sat
c 164	14.8	74.0	972	6	ABU92762	AbI92762 G proteIn	c 237	14.8	74.0	35498	4	ABL20316	ABl20316 Drosophill
c 165	14.8	74.0	1013	6	ABU92760	AbI92760 G proteIn	c 238	14.8	74.0	52216	6	AAH28355	AAH28355 Nucleoicd
c 166	14.8	74.0	1133	7	AAU60086	AaI60086 M. annula	c 239	14.8	74.0	52216	6	ABL50307	ABl50307 Human mus
c 167	14.8	74.0	1165	2	AAV08188	AAv08188 Rabbit in	c 240	14.8	74.0	110000	6	ABS55320_1	ABs55320_1 ContInnation (2 of
c 168	14.8	74.0	1165	2	AAV08188	AAv08188 Rabbit in	c 241	14.8	74.0	110000	6	ABX08336_03	ABx08336_03 ContInnation (4 of
c 169	14.8	74.0	1165	2	AAV63232	AAv63232 Rabbit in	c 242	14.8	74.0	118584	9	ADC87623	ADc87623 Human GPC

C 243	14.8	74.0	126512	6	ABN63429	Abn63429 Human ttx	316	14.2	71.0	165	4	AAK20799	Aak20799 Human bra
244	14.8	74.0	147724	6	ABK63566	Abk63566 Human cdn	317	14.2	71.0	165	4	ABE46714	Abe46714 Human liv
245	14.8	74.0	173810	6	ABN635752	Abn635752 Mouse chr	318	14.2	71.0	165	6	ABE21216	Abe21216 Human gen
246	14.8	74.0	178896	6	ABO68146	Abog68146 Human otc	319	14.2	71.0	170	6	ABX92109	Abx92109 Lung spec
C 247	14.8	74.0	199377	9	ADC35071	Adc35071 Mouse gen	320	14.2	71.0	209	7	ABX43080	Abx43080 Bovine ES
C 248	14.8	74.0	335199	9	ADC24703	Adc24703 Human w11	321	14.2	71.0	218	4	AAH34647	Aah34647 Human col
249	14.8	74.0	341511	6	ABE55200	Abes55200 Genomic D	322	14.2	71.0	230	2	AAAT23211	Aat23211 Human gen
250	14.8	74.0	349999	9	ADC87010	Adc87010 Human GPC	323	14.2	71.0	231	3	AAAC28393	Aac28393 Human sec
C 251	14.4	72.0	3256	6	ABL93045	Ab193045 Rat metae	324	14.2	71.0	234	3	AA667889	Aa667889 Rats rtd
C 252	14.4	72.0	392	8	ACH17198	Ach17198 Human adu	325	14.2	71.0	247	4	AA122630	Aa122630 Probe #12
C 253	14.4	72.0	402	4	AA189035	Aa189035 Human pol	326	14.2	71.0	247	4	ABA67708	Ab67708 Human toe
254	14.4	72.0	540	6	ABN68624	Abn68624 Streptoco	327	14.2	71.0	247	4	AA147924	Aa147924 Probe #16
255	14.4	72.0	576	4	AA523940	Aas23940 Human ova	328	14.2	71.0	247	4	ABA47921	Ab47921 Human bre
256	14.4	72.0	576	5	AAH82498	Aah82498 Human ova	329	14.2	71.0	247	4	ABA34769	Ab34769 Probe #13
257	14.4	72.0	582	4	AA574374	Aas74374 Human ova	330	14.2	71.0	247	4	AAK41863	Aak41863 Human bon
258	14.4	72.0	592	5	AAH82967	Aah82967 Human ova	331	14.2	71.0	247	4	AAK16122	Aak16122 Human bra
C 259	14.4	72.0	621	9	ADD34944	Add34944 Mouse mlt	332	14.2	71.0	247	4	ABE41464	Ab41464 Human liv
C 260	14.4	72.0	637	9	ADB54998	Adb54998 Toxicity-	333	14.2	71.0	247	5	AA108301	Aa108301 Probe #82
C 261	14.4	72.0	1155	6	AA598873	Aas98873 Breast tu	334	14.2	71.0	247	6	ABE15890	Ab15890 Human gen
C 262	14.4	72.0	1155	6	ABO86166	Abog86166 Novel hum	335	14.2	71.0	255	5	AAH82259	Aah82259 Rat diffe
C 263	14.4	72.0	1155	9	ADA11408	Ada11408 Human bre	336	14.2	71.0	283	5	ABV47290	Abv47290 Human pro
C 264	14.4	72.0	1155	9	ADC15381	Adc15381 Human bre	337	14.2	71.0	298	7	ABX42500	Abx42500 Bovine ES
C 265	14.4	72.0	1365	7	ADA70035	Ada70035 Rice gene	338	14.2	71.0	300	2	AAV75297	Aav75297 Staphyloc
C 266	14.4	72.0	1452	3	AA5C1127	Aac51127 Arabidops	339	14.2	71.0	329	3	AACT71145	Aac71145 Single nu
C 267	14.4	72.0	1554	7	ACF72121	Acf72121 Photorhab	340	14.2	71.0	329	3	AACT71148	Aac71148 Single nu
C 268	14.4	72.0	1661	3	AA562273	Aac562273 Arabidops	341	14.2	71.0	329	3	AACT71148	Aac71148 Single nu
C 269	14.4	72.0	1722	7	ACA31375	Ac31375 Prokaryot	342	14.2	71.0	329	3	AACT71148	Aac71148 Single nu
C 270	14.4	72.0	1953	4	AAE61012	Aae61012 P. putida	343	14.2	71.0	330	4	AA527147	Aa527147 CDNA enco
C 271	14.4	72.0	2000	7	ADA73213	Ada73213 Rice gene	344	14.2	71.0	330	4	AA530225	Aas30225 DNA enco
C 272	14.4	72.0	2295	7	ACA36218	Ac36218 Prokaryot	345	14.2	71.0	330	4	AA535560	Aas35560 Human car
C 273	14.4	72.0	2433	2	AAO89728	Aao89728 Plasmodiu	346	14.2	71.0	330	7	ACD01480	Acd01480 Human pol
C 274	14.4	72.0	2437	7	ABT32215	Abt32215 Human neu	347	14.2	71.0	330	9	ADB93325	Adb93325 Human cdn
C 275	14.4	72.0	2492	4	ABE23842	Ab123842 Drosophill	348	14.2	71.0	330	9	ADB45639	Ad45639 Human car
C 276	14.4	72.0	2516	4	ABE25452	Ab125452 Drosophill	349	14.2	71.0	343	5	ABE11960	Ab11960 Human ner
C 277	14.4	72.0	2554	3	AAZ29341	Aaz29341 Tomato be	350	14.2	71.0	343	5	ABE11961	Ab11961 Human ner
C 278	14.4	72.0	2663	2	AAO89728	Aao89728 Plasmodiu	351	14.2	71.0	348	7	ABE240462	Ab240462 N. gonorr
C 279	14.4	72.0	2871	4	ABE20186	Ab120186 Drosophill	352	14.2	71.0	361	5	ABE11559	Ab11559 Human ner
C 280	14.4	72.0	2871	4	ABE22812	Ab122812 Drosophill	353	14.2	71.0	379	6	ABN92985	Abn92985 Staphyloc
C 281	14.4	72.0	2937	4	ABE21888	Ab121888 Drosophill	354	14.2	71.0	379	6	ABN22823	Abn22823 Human ORF
C 282	14.4	72.0	3005	7	AAK99594	Aak99594 MDDT re1a	355	14.2	71.0	381	7	ACA48990	Ac48990 Prokaryot
C 283	14.4	72.0	4062	2	AAAT13347	Aat13347 C3G prote	356	14.2	71.0	382	7	ABE46334	Ab46334 Bovine ES
C 284	14.4	72.0	4062	6	AA594896	Aas94896 Human DNA	357	14.2	71.0	384	7	ACA50945	Aca50945 Prokaryot
C 285	14.4	72.0	9235	2	AAAX13131	Aax13131 Enterococ	358	14.2	71.0	384	7	ACA51451	Aca51451 Prokaryot
C 286	14.4	72.0	9235	6	AB598926	Ab598926 Enterococ	359	14.2	71.0	395	2	AAE83730	Aae83730 DNA enco
C 287	14.4	72.0	11337	4	AAE36286	Aa136286 Human mus	360	14.2	71.0	401	4	AA115814	Aa115814 Probe #57
C 288	14.4	72.0	11337	7	ABX59274	Abx59274 CDNA enco	361	14.2	71.0	401	4	ABE58027	Ab58027 Human gen
C 289	14.4	72.0	12776	4	ABE11894	Ab11894 Drosophill	362	14.2	71.0	401	4	AA137628	Aa137628 Probe #63
C 290	14.4	72.0	12878	4	ABE13212	Ab113212 Drosophill	363	14.2	71.0	401	4	ABE27288	Ab27288 Probe #57
C 291	14.4	72.0	16656	2	AAAT30787	Aat30787 Alphasvitu	364	14.2	71.0	401	4	AAK31752	Aak31752 Human bon
C 292	14.4	72.0	24079	7	AAK71153	Aak71153 Human imm	365	14.2	71.0	401	4	AAK06101	Aak06101 Human bra
C 293	14.4	72.0	24079	7	ABE27330	Ab27330 Human sec	366	14.2	71.0	401	4	ABE31436	Ab31436 Human liv
C 294	14.4	72.0	24079	7	ABE273732	Ab273732 Seceted	367	14.2	71.0	401	6	ABE06508	Ab606508 Human gen
C 295	14.4	72.0	24079	7	ADA98437	Ada98437 Human sec	368	14.2	71.0	433	7	ABX39256	Abx39256 Bovine ES
C 296	14.4	72.0	24079	9	ADC20606	Adc20606 Human sec	369	14.2	71.0	437	8	ACH44707	Ach44707 Human toe
297	14.4	72.0	50000	9	ADC60733	Adc60733 Human sli	370	14.2	71.0	465	3	AAAC37114	Aac37114 Arabidops
298	14.4	72.0	50341	3	AAV26274	Aav26274 DNA seque	371	14.2	71.0	466	4	AA110463	Aa110463 Probe #39
299	14.4	72.0	50341	3	AAZ29519	Aaz29519 US snutcl	372	14.2	71.0	466	4	ABE52109	Ab52109 Human toe
300	14.4	72.0	52297	2	AAAT51411	Aat51411 Mycobacte	373	14.2	71.0	466	4	AA131716	Aa131716 Probe #40
301	14.4	72.0	52298	2	AAQ47357	Aaq47357 U5 mycoba	374	14.2	71.0	466	4	ABE21919	Ab21919 Probe #38
C 302	14.4	72.0	67571	7	AAAD53224_5	Continuation (6 of	375	14.2	71.0	466	4	AAK25834	Aak25834 Human bon
C 303	14.4	72.0	75798	7	ABX77212	Abx77212 Mouse urt	376	14.2	71.0	466	4	AAK00391	Aak00391 Human bra
304	14.4	72.0	110000	7	ACF67367_53	Continuation (54 of	377	14.2	71.0	466	4	ABE25422	Ab25422 Human liv
305	14.4	72.0	110000	7	ACF65367_5	Continuation (6 of	378	14.2	71.0	466	5	AA100399	Aa100399 Probe #39
C 306	14.2	71.0	23	9	ADB88615	Adb88615 Frizled-	379	14.2	71.0	466	6	ABE00415	Ab000415 Human gen
307	14.2	71.0	33	6	AAAL47523	Aal47523 Human rib	380	14.2	71.0	473	6	ABE56963	Ab56963 Human toe
308	14.2	71.0	36	3	AAZ290660	Aaz290660 Nostoc Ns	381	14.2	71.0	473	6	ABE05259	Ab050259 Human gen
C 309	14.2	71.0	65	6	ABN51703	Abn51703 Mouse sp1	382	14.2	71.0	477	8	ACH24857	Ach24857 Human adu
C 310	14.2	71.0	160	6	ABE69525	Ab69525 Human foe	383	14.2	71.0	482	8	ACH46337	Ach46337 Human inf
C 311	14.2	71.0	160	6	ABE17760	Ab17760 Human gen	384	14.2	71.0	484	4	AA113401	Aa113401 Probe #33
312	14.2	71.0	165	4	ABE42373	Ab42373 Human foe	385	14.2	71.0	484	4	ABE55099	Ab55099 Human toe
313	14.2	71.0	165	4	AA152785	Aa152785 Probe #21	386	14.2	71.0	484	4	AA134756	Aa134756 Probe #34
314	14.2	71.0	165	4	ABE38182	Ab38182 Probe #16	387	14.2	71.0	484	4	ABE44645	Ab444645 Human bra
315	14.2	71.0	165	4	AAK46949	Aak46949 Human bon	388	14.2	71.0	484	4	ABE24688	Ab24688 Probe #33

389	14.2	71.0	484	4	AAK28813	AAk28813	Human	bon	C 462	14.2	71.0	1452	6	AB214597	Ab214597	Arabidops
390	14.2	71.0	484	4	AAKO3362	AAKO3362	Human	bra	C 463	14.2	71.0	1470	8	AB213991	Ab213991	Arabidops
391	14.2	71.0	484	4	AB284422	AB284422	Human	liv	C 464	14.2	71.0	1477	8	ADA09909	ADa09909	Human
392	14.2	71.0	484	5	AAI03289	AAi03289	Probe	#32	C 465	14.2	71.0	1481	5	AA880066	AA880066	DNA
393	14.2	71.0	484	6	AB503263	AB503263	Human	gen	C 466	14.2	71.0	1536	6	ABN70708	ABN70708	Strepto
394	14.2	71.0	486	4	AB5457502	ABa57502	Human	foe	C 467	14.2	71.0	1574	6	ABL51009	ABL51009	Human
395	14.2	71.0	486	4	AAI37055	AAi37055	Probe	#57	C 468	14.2	71.0	1594	8	ACDB07895	ACD07895	CDNA
396	14.2	71.0	486	4	AB26990	ABa26990	Probe	#54	C 469	14.2	71.0	1591	4	ABD69393	ABD69393	C. neofo
397	14.2	71.0	486	4	AAK31148	AAK31148	Human	bon	C 470	14.2	71.0	1592	3	AAK73871	AAK73871	Human
398	14.2	71.0	486	4	AAK05545	AAK05545	Human	bra	C 471	14.2	71.0	1604	4	AAK81317	AAK81317	Human
399	14.2	71.0	486	4	AB530828	AB530828	Human	liv	C 472	14.2	71.0	1626	6	ABN70294	ABN70294	Strepto
400	14.2	71.0	486	6	AB505900	AB505900	Human	gen	C 473	14.2	71.0	1626	6	ABN66277	ABN66277	Strepto
401	14.2	71.0	493	4	AAH11545	AAH11545	Human	CDN	C 474	14.2	71.0	1630	3	AAK54015	AAK54015	Arabidops
402	14.2	71.0	525	4	AAHS3240	AAHS3240	S. epider		C 475	14.2	71.0	1635	3	AAK53245	AAK53245	Arabidops
403	14.2	71.0	548	4	ABA59832	ABa59832	Human	foe	C 476	14.2	71.0	1641	3	AAK59785	AAK59785	Human
404	14.2	71.0	548	4	AAI39704	AAi39704	Probe	#83	C 477	14.2	71.0	1642	4	ABL10295	ABL10295	Drosophil
405	14.2	71.0	548	4	ABA28303	ABa28303	Probe	#67	C 478	14.2	71.0	1651	8	ACD98388	ACD98388	Human
406	14.2	71.0	548	4	AAK33981	AAK33981	Human	bon	C 479	14.2	71.0	1651	8	ADD30267	ADD30267	Plant
407	14.2	71.0	548	4	AAK08105	AAK08105	Human	bra	C 480	14.2	71.0	1653	7	ACA339754	ACA339754	Prokaryot
408	14.2	71.0	548	4	AB533784	AB533784	Human	liv	C 481	14.2	71.0	1661	6	ABN68861	ABN68861	Strepto
409	14.2	71.0	548	6	AB508782	AB508782	Human	gen	C 482	14.2	71.0	1707	6	ABN91468	ABN91468	Staphyloc
410	14.2	71.0	555	9	ADC91108	ADc91108	E. fecali		C 483	14.2	71.0	1725	7	ACA43064	ACA43064	Prokaryot
411	14.2	71.0	562	9	ADD33014	ADd33014	Human	mit	C 484	14.2	71.0	1790	7	AAK63505	AAK63505	Human
412	14.2	71.0	571	5	ABV35222	ABv35222	Human	pro	C 485	14.2	71.0	1790	10	ADE76985	ADE76985	Human
413	14.2	71.0	586	6	ABN66037	ABN66037	Human	can	C 486	14.2	71.0	1816	4	AAK25489	AAK25489	Nucleotid
414	14.2	71.0	592	4	AAK79691	AAK79691	Human	imm	C 487	14.2	71.0	1820	7	ADA70905	ADA70905	Rice
415	14.2	71.0	600	5	ABV14127	ABv14127	Human	pro	C 488	14.2	71.0	1841	7	AB236174	AB236174	Human
416	14.2	71.0	629	3	AAK09327	AAK09327	Fusarium		C 489	14.2	71.0	1843	2	AAO14798	AAO14798	Ro/SSA
417	14.2	71.0	658	4	ABL15187	ABl15187	Drosophil		C 490	14.2	71.0	1845	3	AAZ38967	AAZ38967	Escherich
418	14.2	71.0	657	4	ABL22877	ABl22877	Drosophil		C 491	14.2	71.0	1875	6	ABL59530	ABL59530	Human
419	14.2	71.0	693	7	AB272013	ABz72013	Human	NAD	C 492	14.2	71.0	1883	6	ADD18526	ADD18526	Human
420	14.2	71.0	700	4	ABL03751	ABl03751	Drosophil		C 493	14.2	71.0	1890	6	ABK84052	ABK84052	Human
421	14.2	71.0	721	2	AAAT35105	AAAT35105	Down-regu		C 494	14.2	71.0	1899	7	ACA21814	ACA21814	Prokaryot
422	14.2	71.0	743	8	ABL51017	ABl51017	Human	EDA	C 495	14.2	71.0	1902	6	ABO70018	ABO70018	Listeria
423	14.2	71.0	743	8	ACD07993	ACd07993	DNA	encod	C 496	14.2	71.0	1903	6	ABO68241	ABO68241	Listeria
424	14.2	71.0	744	6	ABK35418	ABK35418	Human	CDN	C 497	14.2	71.0	1903	4	AAH16546	AAH16546	Human
425	14.2	71.0	771	4	AAK53200	AAK53200	Arabidops		C 498	14.2	71.0	1908	7	ACA26970	ACA26970	Prokaryot
426	14.2	71.0	776	6	ABN66321	ABn66321	Streptoco		C 499	14.2	71.0	1953	3	AAA09499	AAA09499	Plant-opt
427	14.2	71.0	776	4	AAI95828	AAi95828	Human	neu	C 500	14.2	71.0	1962	3	AAZ51637	AAZ51637	Delta-ent
428	14.2	71.0	789	3	AAK53161	AAK53161	Arabidops		C 501	14.2	71.0	1962	7	ABX95181	ABX95181	DNA
429	14.2	71.0	801	2	AAH07525	AAH07525	Human	CDN	C 502	14.2	71.0	1962	9	AAAD61785	AAAD61785	Bacillus
430	14.2	71.0	807	4	AAK39942	AAK39942	Gaetlic	C	C 503	14.2	71.0	1989	9	AAZ51638	AAZ51638	Bacillus
431	14.2	71.0	830	2	AA216633	AA216633	Human	gen	C 504	14.2	71.0	1989	6	AB565041	AB565041	Invertebr
432	14.2	71.0	870	7	ACA39785	ACA39785	Prokaryot		C 505	14.2	71.0	1989	7	ABX95182	ABX95182	DNA
433	14.2	71.0	896	5	ABAI5574	ABAI5574	Human	ner	C 506	14.2	71.0	1989	9	AAAD61786	AAAD61786	Bacillus
434	14.2	71.0	915	7	ACA51998	ACA51998	Prokaryot		C 507	14.2	71.0	2000	2	AAK81529	AAK81529	Plantized
435	14.2	71.0	917	3	AAK36836	AAK36836	Arabidops		C 508	14.2	71.0	2000	7	ADA73099	ADA73099	Rice
436	14.2	71.0	919	6	ABUS8803	ABi58803	Fungal	me	C 509	14.2	71.0	2000	7	ADA72245	ADA72245	Rice
437	14.2	71.0	930	7	ACA39461	ACA39461	Prokaryot		C 510	14.2	71.0	2000	7	ADA72245	ADA72245	Rice
438	14.2	71.0	1025	6	AAAD1879	AAAD1879	Lactobacil		C 511	14.2	71.0	2002	2	AAV43605	AAV43605	Human
439	14.2	71.0	1059	6	ABL40697	ABl40697	Human	big	C 512	14.2	71.0	2020	2	AAQ53209	AAQ53209	Cycilin D3
440	14.2	71.0	1095	4	ABL22477	ABl22477	Drosophil		C 513	14.2	71.0	2022	2	AAO31877	AAO31877	DNA
441	14.2	71.0	1146	7	ACA32818	ACA32818	Prokaryot		C 514	14.2	71.0	2050	2	AAK81530	AAK81530	DNA
442	14.2	71.0	1148	9	ADD30958	ADD30958	Plant	yle	C 515	14.2	71.0	2109	3	AAK98855	AAK98855	Human
443	14.2	71.0	1176	6	ABL51020	ABl51020	Human	EDA	C 516	14.2	71.0	2110	2	ADB62589	ADB62589	Human
444	14.2	71.0	1176	8	ACD07906	ACd07906	DNA	encod	C 517	14.2	71.0	2152	2	AAO70732	AAO70732	TATRA-bind
445	14.2	71.0	1192	3	AAK36086	AAK36086	Arabidops		C 518	14.2	71.0	2152	2	AAAT42218	AAAT42218	Human
446	14.2	71.0	1213	4	AAK63802	AAK63802	Human	sec	C 519	14.2	71.0	2152	2	AAAT79594	AAAT79594	TATRA-bind
447	14.2	71.0	1214	5	AAK29071	AAK29071	CDNA	encod	C 520	14.2	71.0	2200	3	AAK77823	AAK77823	Human
448	14.2	71.0	1214	6	AB568211	AB568211	CDNA	encod	C 521	14.2	71.0	2211	5	AAK91228	AAK91228	Human
449	14.2	71.0	1214	6	ADC5205	ADC5205	Human	CDN	C 522	14.2	71.0	2215	7	ACA45338	ACA45338	Prokaryot
450	14.2	71.0	1248	7	AAAT3335	AAAT3335	American		C 523	14.2	71.0	2232	5	AA594081	AA594081	DNA
451	14.2	71.0	1248	7	ACA848457	ACA848457	Prokaryot		C 524	14.2	71.0	2259	2	AAAT30686	AAAT30686	Kaposi's
452	14.2	71.0	1266	7	ABX06878	ABx06878	S. pneumo		C 525	14.2	71.0	2259	2	AAAT16816	AAAT16816	Kaposi's
453	14.2	71.0	1281	7	ACA39944	ACA39944	Prokaryot		C 526	14.2	71.0	2274	7	ACA43678	ACA43678	Prokaryot
454	14.2	71.0	1287	7	ACA31014	ACA31014	Prokaryot		C 527	14.2	71.0	2305	7	ADB89728	ADB89728	Staphyloc
455	14.2	71.0	1290	5	AA529143	AAa29143	CDNA	encod	C 528	14.2	71.0	2321	9	ADB56281	ADB56281	Human
456	14.2	71.0	1290	5	AB568283	AB568283	CDNA	encod	C 529	14.2	71.0	2321	9	ADB56281	ADB56281	Human
457	14.2	71.0	1290	9	ADC52277	ADC52277	Human	CDN	C 530	14.2	71.0	2344	2	AAO33394	AAO33394	R496L
458	14.2	71.0	1305	3	AAK54212	AAK54212	Arabidops		C 531	14.2	71.0	2344	8	AAAT95066	AAAT95066	CDNA
459	14.2	71.0	1312	4	ABL08777	ABl08777	Drosophil		C 532	14.2	71.0	2344	8	ACD28698	ACD28698	CDNA
460	14.2	71.0	1365	5	AAK40275	AAK40275	Arabidops		C 533	14.2	71.0	2346	2	AAAT95068	AAAT95068	CDNA
461	14.2	71.0	1452	5	AA571632	AA571632	DNA	encod	C 534	14.2	71.0	2347	2	AAO33390	AAO33390	ASM



c 681	14.2	71.0	3265	7	ACC91122	ACC91122 Human sec	c 754	14.2	71.0	3265	7	ACA98482	ACA98482 Human PRO
c 682	14.2	71.0	3265	7	ACC88864	ACC88864 Human sec	c 755	14.2	71.0	3265	7	ABX17152	ABX17152 Human PRO
c 683	14.2	71.0	3265	7	ACD07061	ACD07061 Human PRO	c 756	14.2	71.0	3265	7	ABX16714	ABX16714 Human CDN
c 684	14.2	71.0	3265	7	ACA67512	ACA67512 Human PRO	c 757	14.2	71.0	3265	8	ACA68007	ACA68007 Novel hum
c 685	14.2	71.0	3265	7	ACC81567	ACC81567 Human sec	c 758	14.2	71.0	3265	8	ACA63407	ACA63407 CDNA enco
c 686	14.2	71.0	3265	7	ACA91284	ACA91284 CDNA enco	c 759	14.2	71.0	3265	8	ACA97655	ACA97655 Human PRO
c 687	14.2	71.0	3265	7	ACC89171	ACC89171 Human sec	c 760	14.2	71.0	3265	8	ACA99104	ACA99104 Novel hum
c 688	14.2	71.0	3265	7	ACC86527	ACC86527 Human sec	c 761	14.2	71.0	3265	8	ACC991736	ACC991736 Human sec
c 689	14.2	71.0	3265	7	ACC89785	ACC89785 Human sec	c 762	14.2	71.0	3265	8	ACD11147	ACD11147 Novel hum
c 690	14.2	71.0	3265	7	ACC92964	ACC92964 Human sec	c 763	14.2	71.0	3265	8	ACD14997	ACD14997 Human sec
c 691	14.2	71.0	3265	7	ABX80869	ABX80869 Human sec	c 764	14.2	71.0	3265	8	ACA88456	ACA88456 Human sec
c 692	14.2	71.0	3265	7	ACA72592	ACA72592 Human PRO	c 765	14.2	71.0	3265	8	ACD81963	ACD81963 CDNA enco
c 693	14.2	71.0	3265	7	ACA89110	ACA89110 Human sec	c 766	14.2	71.0	3265	8	ACD11761	ACD11761 Human sec
c 694	14.2	71.0	3265	7	ACA69886	ACA69886 Human sec	c 767	14.2	71.0	3265	8	ACC95890	ACC95890 Human sec
c 695	14.2	71.0	3265	7	ACA96989	ACA96989 Novel hum	c 768	14.2	71.0	3265	8	ACF16453	ACF16453 Human sec
c 696	14.2	71.0	3265	7	ACA90985	ACA90985 Novel hum	c 769	14.2	71.0	3265	8	ACF02571	ACF02571 Human sec
c 697	14.2	71.0	3265	7	ACA70767	ACA70767 Human sec	c 770	14.2	71.0	3265	8	ACF02878	ACF02878 Human sec
c 698	14.2	71.0	3265	7	ACA95277	ACA95277 Novel hum	c 771	14.2	71.0	3265	8	ACF21465	ACF21465 Human sec
c 699	14.2	71.0	3265	7	ACD44378	ACD44378 CDNA enco	c 772	14.2	71.0	3265	8	ACF10149	ACF10149 Human sec
c 700	14.2	71.0	3265	7	ACC86220	ACC86220 Human sec	c 773	14.2	71.0	3265	8	ACF78042	ACF78042 Human sec
c 701	14.2	71.0	3265	7	ACD45183	ACD45183 Human sec	c 774	14.2	71.0	3265	8	ACD46747	ACD46747 Human sec
c 702	14.2	71.0	3265	7	ACC90092	ACC90092 Human sec	c 775	14.2	71.0	3265	8	ACD49510	ACD49510 Human sec
c 703	14.2	71.0	3265	7	ACD12700	ACD12700 Human sec	c 776	14.2	71.0	3265	8	ACF28277	ACF28277 Human sec
c 704	14.2	71.0	3265	7	ACF19930	ACF19930 Human sec	c 777	14.2	71.0	3265	8	ACD88967	ACD88967 Human sec
c 705	14.2	71.0	3265	7	ABX76874	ABX76874 Human PRO	c 778	14.2	71.0	3265	8	ACD84362	ACD84362 Human PRO
c 706	14.2	71.0	3265	7	ACA73206	ACA73206 Novel hum	c 779	14.2	71.0	3265	8	ACD99136	ACD99136 CDNA enco
c 707	14.2	71.0	3265	7	ACA68749	ACA68749 Novel hum	c 780	14.2	71.0	3265	8	ADA78009	ADA78009 Human sec
c 708	14.2	71.0	3265	7	ACA74593	ACA74593 CDNA enco	c 781	14.2	71.0	3265	8	ACF48878	ACF48878 Human sec
c 709	14.2	71.0	3265	7	ACA70460	ACA70460 Human sec	c 782	14.2	71.0	3265	8	ACD09198	ACD09198 Human sec
c 710	14.2	71.0	3265	7	ACD14646	ACD14646 Human PRO	c 783	14.2	71.0	3265	8	ACF11991	ACF11991 Human sec
c 711	14.2	71.0	3265	7	ACA93731	ACA93731 Human CDN	c 784	14.2	71.0	3265	8	ACF41225	ACF41225 Human sec
c 712	14.2	71.0	3265	7	ACA68318	ACA68318 Novel hum	c 785	14.2	71.0	3265	8	ACF15839	ACF15839 Human sec
c 713	14.2	71.0	3265	7	ABX98783	ABX98783 Novel hum	c 786	14.2	71.0	3265	8	ACF16146	ACF16146 Human sec
c 714	14.2	71.0	3265	7	ACA67305	ACA67305 CDNA enco	c 787	14.2	71.0	3265	8	ADB17126	ADB17126 Human CDN
c 715	14.2	71.0	3265	7	ACC81260	ACC81260 Human sec	c 788	14.2	71.0	3265	8	ACD31973	ACD31973 Human sec
c 716	14.2	71.0	3265	7	ACA95584	ACA95584 Novel hum	c 789	14.2	71.0	3265	8	ACF18781	ACF18781 Human sec
c 717	14.2	71.0	3265	7	ACD04502	ACD04502 Novel hum	c 790	14.2	71.0	3265	8	ACF09228	ACF09228 Human sec
c 718	14.2	71.0	3265	7	ACCF7943	ACCF7943 Human sec	c 791	14.2	71.0	3265	8	ACF78349	ACF78349 Human sec
c 719	14.2	71.0	3265	7	ACCF12605	ACCF12605 Human sec	c 792	14.2	71.0	3265	8	ACF51948	ACF51948 Human sec
c 720	14.2	71.0	3265	7	ACH66278	ACH66278 Novel hum	c 793	14.2	71.0	3265	8	ACF26435	ACF26435 Human sec
c 721	14.2	71.0	3265	7	ABX79549	ABX79549 Human sec	c 794	14.2	71.0	3265	8	ACF24228	ACF24228 Human sec
c 722	14.2	71.0	3265	7	ACA96320	ACA96320 Human PRO	c 795	14.2	71.0	3265	8	ACF63539	ACF63539 Human sec
c 723	14.2	71.0	3265	7	ACA65094	ACA65094 Human PRO	c 796	14.2	71.0	3265	8	ACF50413	ACF50413 Human sec
c 724	14.2	71.0	3265	7	ACA73820	ACA73820 Human PRO	c 797	14.2	71.0	3265	8	ACH07884	ACH07884 Human sec
c 725	14.2	71.0	3265	7	ACA74232	ACA74232 Novel hum	c 798	14.2	71.0	3265	8	ACF13690	ACF13690 Human sec
c 726	14.2	71.0	3265	7	ACA96627	ACA96627 Human PRO	c 799	14.2	71.0	3265	8	ACD41616	ACD41616 Human sec
c 727	14.2	71.0	3265	7	ACD10733	ACD10733 CDNA enco	c 800	14.2	71.0	3265	8	ADA37889	ADA37889 Human CDN
c 728	14.2	71.0	3265	7	ACC91429	ACC91429 Human sec	ALIGNMENTS						
c 729	14.2	71.0	3265	7	ACA93570	ACA93570 Novel hum							
c 730	14.2	71.0	3265	7	ACD02764	ACD02764 CDNA enco							
c 731	14.2	71.0	3265	7	ACC87329	ACC87329 Human sec							
c 732	14.2	71.0	3265	7	ACC85913	ACC85913 Human sec							
c 733	14.2	71.0	3265	7	ABX81252	ABX81252 Novel hum							
c 734	14.2	71.0	3265	7	ACA65401	ACA65401 Human PRO							
c 735	14.2	71.0	3265	7	ACA94218	ACA94218 Human sec							
c 736	14.2	71.0	3265	7	ACA97962	ACA97962 Human PRO							
c 737	14.2	71.0	3265	7	ACA91464	ACA91464 Novel hum							
c 738	14.2	71.0	3265	7	ACA90678	ACA90678 Novel hum							
c 739	14.2	71.0	3265	7	ACD16225	ACD16225 Human sec							
c 740	14.2	71.0	3265	7	ACD17386	ACD17386 Human sec							
c 741	14.2	71.0	3265	7	ACC92043	ACC92043 Human sec							
c 742	14.2	71.0	3265	7	ACD02332	ACD02332 Novel hum							
c 743	14.2	71.0	3265	7	ACA74900	ACA74900 CDNA enco							
c 744	14.2	71.0	3265	7	ACA91771	ACA91771 Human PRO							
c 745	14.2	71.0	3265	7	ACA89323	ACA89323 Novel hum							
c 746	14.2	71.0	3265	7	ACA71415	ACA71415 Human sec							
c 747	14.2	71.0	3265	7	ACC90815	ACC90815 Human sec							
c 748	14.2	71.0	3265	7	ACA65825	ACA65825 CDNA enco							
c 749	14.2	71.0	3265	7	ACA68960	ACA68960 Novel hum							
c 750	14.2	71.0	3265	7	ACA93068	ACA93068 Novel hum							
c 751	14.2	71.0	3265	7	ACA94970	ACA94970 CDNA enco							
c 752	14.2	71.0	3265	7	ACD16532	ACD16532 Human sec							
c 753	14.2	71.0	3265	7	ACD15611	ACD15611 Human sec							

RESULT 1												
AAAD21384												
ID	AAAD21384	strand:	DNA;	20	BP.							
XX	AAAD21384;											
XX	28-JUN-2002	(first entry)										
DE	Antisense oligo,	HYB 963,	directed	against	human XPA	gene.						
XX	Human;	cytotoxin;	cancer;	transcription	coupled	repair;	TCR;					
KW	nucleotide	excision	repair;	NER;	antisense;	cytostatic;						
KM	xeroderma	pigmentosum	group	A;	XPA;	ss.						
FT	key	Location/Qualifiers										
FT	modified_base	1..20										
FT	/*tag=	a										
FT	/mod_base=	OTHER										

RESULT 1  
AAD21384  
ID AAD21384 strand; DNA; 20 BP.

XX AAD21384;

XX 28-JAN-2002 (first entry)

DE Antisense oligo, HYB 963, directed against human XPA gene.

XX Human; cytochrome; cancer; transcription coupled repair; TCR;

KW nucleotide excision repair; NER; antisense; cytoskeletal;

XX Xeroderma pigmentosum group A; XPA; ss.

XX Homo sapiens.

OS Synthetic.

FT key Location/Qualifiers  
FT modified\_base 1..20  
FT /\*tag= a  
FT /\_mod\_base= OTHER

/note= "Phosphorothioate backbone"

PT WO200174346-A2.  
XX 11-OCT-2001.  
XX 03-APR-2001; 2001WO-US010800.  
XX 03-APR-2000; 2000US-0194343P.  
XX (HYBR-) HYBRIDON INC.  
XX Agrawal S, Kandimalia ER, Bregman DB, Mani S, Lu Y;  
XX MPI; 2001-662947/76.  
XX  
XX Increasing sensitivity of cancer cells to a cytotoxin or oxidizing agent  
XX useful for therapy comprises contacting them with oligonucleotides  
XX complementary to transcription coupled repair or nucleotide excision  
XX repair genes.  
XX  
XX Claim 12; Page 18; 58pp; English.  
XX  
XX The present invention relates to a method for potentiating or enhancing  
XX the toxic effect of a cytotoxin or oxidizing agent on a cancer cell,  
XX comprising contacting the cell with an oligonucleotide complementary to a  
XX gene involved in transcription coupled repair (TCR) and nucleotide  
XX excision repair (NER) and with a cytotoxin or oxidizing agent. The  
XX invention is used to sensitize cancer cells to therapeutic agents. The  
XX present sequence is an antisense oligonucleotide directed against  
XX Xeroderma pigmentosum group A (XPA) gene  
XX  
SQ Sequence 20 BP; 4 A; 4 C; 5 G; 7 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
DB 1 GGTCCTACTCATGTTGATG 20

RESULT 2  
ABL84566  
ID ABL84566 standard; cDNA, 486 BP.  
XX  
XX ABL84566;  
XX  
XX 17-MAY-2002 (first entry)  
XX  
XX Human ovarian cancer related cDNA clone SEQ ID NO:7544.  
XX  
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200192581-A2.  
XX  
XX 06-DEC-2001.  
XX  
XX 29-MAY-2001; 2001WO-US017756.  
XX  
XX 26-MAY-2000; 2000US-0207484P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Algate PA, Harlocker SL, Jones R;  
XX  
XX MPI; 2002-122075/16.  
XX  
XX Composition for therapy and diagnosis of ovarian cancer comprising  
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding

PT polypeptide, antibody specific to polypeptide or T cell expressing  
PT polypeptide.  
XX  
XX Claim 1; SEQ ID NO 7544; 489pp; English.  
XX  
XX The present invention describes a composition (I) comprising: carriers  
XX and immunostimulants; and a polypeptide (II) of a ovarian tumour  
XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)  
XX from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,  
XX (III) encoding (II) having a sequence (S2), a T cell population of (II),  
XX or antigen presenting cells that express (II). (I) has cytostatic  
XX activity. An oligonucleotide (IV) that hybridizes to (S1) can be used for  
XX detecting ovarian cancer in a patient's biological sample preferably  
XX serum or ovarian tissue. The method comprises contacting a biological  
XX sample from a patient with (IV), detecting the amount of polynucleotide  
XX hybridizing to (IV) and comparing the amount to a predetermined cutoff  
XX value and thereby detecting ovarian cancer in the patient, where the  
XX amount of polynucleotide hybridizing to (IV) is detected preferably by  
XX polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
XX useful for stimulating and/or expanding T cells specific for an ovarian  
XX tumour protein comprising contacting T cells with (III) or (II). (III) is  
XX useful in design and preparation of ribozyme molecules for inhibiting  
XX expression of the tumour polypeptides and proteins in tumour cells; and  
XX to isolate a full length gene from a suitable library e.g., a tumour cDNA  
XX library using well known techniques  
XX  
SQ Sequence 486 BP; 121 A; 106 C; 70 G; 188 T; 0 U; 1 Other;  
Query Match 100.0%; Score 20; DB 6; Length 486;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
DB 149 GGTCCTACTCATGTTGATG 168

RESULT 3  
ABL79139/c  
ID ABL79139 standard; cDNA, 506 BP.  
XX  
XX ABL79139;  
XX  
XX 17-MAY-2002 (first entry)  
XX  
XX Human ovarian cancer related cDNA clone SEQ ID NO:2117.  
XX  
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200192581-A2.  
XX  
XX 06-DEC-2001.  
XX  
XX 29-MAY-2001; 2001WO-US017756.  
XX  
XX 26-MAY-2000; 2000US-0207484P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Algate PA, Harlocker SL, Jones R;  
XX  
XX MPI; 2002-122075/16.  
XX  
XX Composition for therapy and diagnosis of ovarian cancer comprising  
XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
XX polypeptide, antibody specific to polypeptide or T cell expressing  
XX polypeptide.  
XX  
XX Claim 1; SEQ ID NO 2117; 489pp; English.  
XX  
XX The present invention describes a composition (I) comprising: carriers



CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)  
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,  
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),  
 CC or antigen presenting cells that express (II). (I) has cytostatic  
 CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for  
 CC detecting ovarian cancer in a patient's biological sample preferably  
 CC serum or ovarian tissue. The method comprises contacting a biological  
 CC sample from a patient with (IV), detecting the amount of polynucleotide  
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
 CC value and thereby detecting ovarian cancer in the patient, where the  
 CC amount of polynucleotide hybridising to (IV) is detected preferably by  
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
 CC useful for stimulating and/or expanding T cells specific for an ovarian  
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is  
 CC useful in design and preparation of ribozyme molecules for inhibiting  
 CC expression of the tumour polypeptides and proteins in tumour cells; and  
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
 CC library using well known techniques

SO Sequence 506 BP; 174 A; 95 C; 104 G; 132 T; 0 U; 1 Other;

Query Match 100.0%; Score 20; DB 6; Length 506;  
 Best Local Similarity 100.0%; Pred. NO. 3.6;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCCTACTCATGTTGATG 20  
 |||||  
 DB 246 GGTCCTACTCATGTTGATG 227

RESULT 4  
 ABL79129  
 ID ABL79129 standard; cDNA; 513 BP.

XX ABL79129;

XX 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:2107.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX MO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising  
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
 PT polypeptide, antibody specific to polypeptide or T cell expressing  
 PT polypeptide.

XX Claim 1; SEQ ID NO 2107; 489bp; English.

XX The present invention describes a composition (I) comprising: carriers  
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)  
 CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,  
 CC (III) encoding (II) having a sequence (S2), a T cell population of (II),  
 CC or antigen presenting cells that express (II). (I) has cytostatic  
 CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for

CC detecting ovarian cancer in a patient's biological sample preferably  
 CC serum or ovarian tissue. The method comprises contacting a biological  
 CC sample from a patient with (IV), detecting the amount of polynucleotide  
 CC hybridising to (IV) and comparing the amount to a predetermined cutoff  
 CC value and thereby detecting ovarian cancer in the patient, where the  
 CC amount of polynucleotide hybridising to (IV) is detected preferably by  
 CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is  
 CC useful for stimulating and/or expanding T cells specific for an ovarian  
 CC tumour protein comprising contacting T cells with (III) or (II). (III) is  
 CC useful in design and preparation of ribozyme molecules for inhibiting  
 CC expression of the tumour polypeptides and proteins in tumour cells; and  
 CC to isolate a full length gene from a suitable library e.g., a tumour cDNA  
 CC library using well known techniques

SO Sequence 513 BP; 144 A; 95 C; 100 G; 166 T; 0 U; 8 Other;

Query Match 100.0%; Score 20; DB 6; Length 513;  
 Best Local Similarity 100.0%; Pred. NO. 3.7; Indels 0; Gaps 0;  
 Matches 20; Conservative 0; Mismatches 0;

OY 1 GGTCCTACTCATGTTGATG 20  
 |||||  
 DB 357 GGTCCTACTCATGTTGATG 376

RESULT 5  
 AAS63227/C  
 ID AAS63227 standard; cDNA; 822 BP.

XX AAS63227;

XX 29-JAN-2002 (first entry)

XX Human xeroderma pigmentosum complementation group A (XPA) DNA.

XX DNA mutation-binding protein; nuclease; DNA mismatch; cancer; PCR primer;

XX DNA damage; human xeroderma pigmentosum complementation group; XPF; XPA;

XX XPC; XPE; ERCC4; human MutS homologue 2; hMSH2; MutS; Nuc; MutY; Fpg; ss;

XX Fapy-DNA glycosylase; uracil DNA glycosylase; ung; Tdg; xthA gene; Uvr A;

XX A/G-specific adenine glycosylase; synthetic T4 endonuclease V; T4 endo V;

XX thymine DNA-glycosylase; Uvr B; Uvr C; nth gene; nfo gene; exonuclease;

XX endonuclease.

XX Homo sapiens.

XX MO200173079-A2.

XX 04-OCT-2001.

XX 26-MAR-2001; 2001WO-US009700.

XX 28-MAR-2000; 2000US-0192764P.

XX 29-AUG-2000; 2000US-00650855.

XX (RECG ) UNIT CALIFORNIA.

XX Mc Cutchen- Maloney SL;

XX WPI; 2001-656920/75.

XX P-PSDB; AAU69740.

XX Recombinant chimeric protein, useful for detecting and quantifying DNA

XX mutations, e.g. in disease diagnosis, comprises mutation-binding protein

XX and nuclease.

XX Claim 46; Page 62; 128bp; English.

XX Sequences AAS63224-AAS63305 represent cDNA encoding and PCR primers used



CC proteins are able to bind to the site of the DNA mutation and cut it out  
CC of the molecule. This is useful for early diagnosis of cancer and other  
CC diseases. The proteins used in the invention include human XPF (or  
CC ERCC1), human xeroderma pigmentosum complementation groups A, C and E  
CC (XPA, XPC and XPD), human Muts homologue 2 (hMSH2), *Serratia marcescens*  
CC nuclease (Nuc), *Thermus thermophilus* Muts, *Escherichia coli* Fapy-DNA  
CC glycosylase (Fpg), uracil DNA glycosylase (ung), A/G-specific adenine  
CC glycosylase (Muty), synthetic T4 endonuclease V (T4 endo V), thymine DNA-  
CC glycosylase (TDG), *E. coli* Uvr A, B and C, and *E. coli* endonucleases and  
CC exonucleases  
XX  
SQ Sequence 822 BP; 280 A; 144 C; 225 G; 173 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 4; Length 822;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCATCTCATGTTGATG 20  
DB 743 GGTCATCTCATGTTGATG 724  
RESULT 6  
AAD21394/c  
ID AAD21394 standard; DNA; 1377 BP.  
XX  
AC AAD21394;  
XX  
DT 28-JAN-2002 (first entry)  
XX  
DE Human Xeroderma pigmentosum group A (XPA) DNA.  
XX  
KW Human; cytotoxin; cancer; transcription coupled repair; TCR;  
KW nucleotide excision repair; NER; cytoskeletal; chromosome 9;  
KW Xeroderma pigmentosum group A; XPA; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200174346-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 03-APR-2001; 2001WO-US010800.  
XX  
PR 03-APR-2000; 2000US-0194343P.  
XX  
PA (HYBR-) HYBRIDON INC.  
PI Agrawal S, Kandimalia ER, Bregman DB, Mani S, Lu Y;  
XX WPI; 2001-662947/76.  
DR  
PT Increasing sensitivity of cancer cells to a cytotoxin or oxidizing agent  
PT useful for therapy comprises contacting them with oligonucleotides  
PT complementary to transcription coupled repair or nucleotide excision  
PT repair genes.  
PS Disclosure; Page 48; 58pp; English.  
XX  
CC The present invention relates to a method for potentiating or enhancing  
CC the toxic effect of a cytotoxin or oxidizing agent on a cancer cell,  
CC comprising contacting the cell with an oligonucleotide complementary to a  
CC gene involved in transcription coupled repair (TCR) and nucleotide  
CC excision repair (NER) and with a cytotoxin or oxidizing agent. The  
CC invention is used to sensitize cancer cells to therapeutic agents. The  
CC present sequence is a human Xeroderma pigmentosum group A (XPA) gene  
CC located on chromosome 9  
XX  
SQ Sequence 1377 BP; 458 A; 232 C; 358 G; 329 T; 0 U; 0 Other;  
Query Match 100.0%; Score 20; DB 4; Length 1377;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTTGATG 20  
DB 769 GGTCATCTCATGTTGATG 750  
RESULT 7  
AAS34872/c  
ID AAS34872 standard; cDNA; 1407 BP.  
XX  
AC AAS34872;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE cDNA encoding novel human neoplastic disease associated polypeptide #106.  
XX  
KW Human; neoplastic disease associated polypeptide; cancer; gene therapy;  
KW hyperproliferative disorder; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW neuroprotective; cytoskeletal; anti inflammatory; vasotropic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155163-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001358.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226811P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.



KW Neoplastic disease-associated polypeptide; gene therapy;  
KW hyperproliferative disease; cancer; autoimmune disorder; diabetes;  
KW rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis;  
KW autoimmune thyroiditis; haemolytic anaemia; haematopoietic disorder;  
KW haematologic disorder; anaemia; thrombocytopenia; allergic reaction;  
KW asthma; eczema; inflammatory disorder; lechaemia-reperfusion injury;  
KW inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;  
KW Alzheimer's disease; Parkinson's disease; renal disorder;  
KW acute glomerulonephritis; end-stage renal disease;  
KW cardiovascular disorder; atherosclerosis; myocarditis;  
KW infectious disease; AIDS; cachexia; anorexia; wound healing;  
KW epithelial cell proliferation; Human; ss.  
XX  
XX Homo sapiens.  
XX  
XX US2003082758-A1.  
XX  
XX 01-MAY-2003.  
XX  
XX 22-MAR-2002; 2002US-00103313.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-020515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-022547P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-022679P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-022668P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-023113P.  
PR 08-SEP-2000; 2000US-023114P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232402P.  
PR 14-SEP-2000; 2000US-0232403P.  
PR 14-SEP-2000; 2000US-0232404P.  
PR 14-SEP-2000; 2000US-0232405P.  
PR 14-SEP-2000; 2000US-0232406P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234224P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 PR 17-JAN-2001; 2001US-00764854.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM, Barash SC;  
 PI WPI; 2003-786918/74.  
 DR P-PSDB; ADC46314.  
 XX  
 PT New isolated human neoplastic disease-associated polypeptides and  
 PT polynucleotides, useful for diagnosing, preventing, prognosticating or  
 PT treating medical conditions such as cancer, AIDS, diabetes or Parkinson's  
 PT disease.  
 XX  
 PS Claim 1; SEQ ID NO 116; 302pp; English.  
 XX  
 CC The invention relates to one of 238 disclosed human neoplastic disease-  
 CC associated polypeptides encoded by 171 disclosed cDNA sequences  
 CC (including their domains, epitopes, full-length proteins, allelic variants  
 CC or species homologues). Also included are there encoding nucleic acids, a  
 CC recombinant vector comprising the nucleic acid, a recombinant host cell  
 CC comprising the nucleic acid (expressing the protein), an isolated  
 CC antibody that binds specifically to the isolated polypeptide, preventing,  
 CC treating or ameliorating a medical condition, diagnosing a pathological  
 CC condition or a susceptibility to a pathological condition in a subject,  
 CC identifying a binding partner to the polypeptide, identifying an activator  
 CC in a biological assay, and the gene corresponding to the cDNA sequence.  
 CC The polypeptides, polynucleotides and antibodies are useful for  
 CC detecting, preventing, diagnosing, prognosticating, treating or  
 CC ameliorating medical conditions such as hyperproliferative diseases or  
 CC cancer, autoimmune disorders (e.g. diabetes, rheumatoid arthritis,  
 CC systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis  
 CC or haemolytic anaemia), haematopoietic or haematologic disorders (e.g.  
 CC anaemia or thrombocytopenia), allergic reactions including asthma or  
 CC eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury,  
 CC  
 Query Match 100.0%; Score 20; DB 9; Length 1407;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGTCCATCTCATGTTGATG 20  
 DB 782 GGTCCATCTCATGTTGATG 763  
 RESULT 9  
 ABX35476  
 ID ABX35476 standard; cDNA; 4670 BP.  
 XX  
 AC ABX35476;  
 XX  
 DT 05-FEB-2003 (first entry)  
 XX  
 DE Human gene expression profile polynucleotide SEQ ID NO 587.  
 XX  
 KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;  
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;  
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;  
 KW gene expression; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX

PN WO200274979-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PP 20-MAR-2002; 2002MO-US008456.  
 XX  
 PR 20-MAR-2001; 2001US-0276947P.  
 XX  
 PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
 XX  
 PI Wan J, Mang Y;  
 XX  
 DR WPI; 2002-740862/80.  
 XX  
 PT New gene expression profile generated from primary, endothelial,  
 PT epithelial, and muscle cell types, useful for identifying disease  
 PT pathologies involving alterations of gene expression, e.g. cancer.  
 XX  
 PS Disclosure; Page 716-718; 850pp; English.  
 XX  
 CC The invention relates to a gene expression profile comprising one or more  
 CC genes (ABX34889-ABX35692) and generated from a cell type. The cell type  
 CC is a coronary artery endothelium, umbilical artery or vein endothelium,  
 CC aortic endothelium, dermal microvascular endothelium, pulmonary artery  
 CC endothelium, myometrium microvascular endothelium, keratinocyte  
 CC epithelium, bronchial epithelium, mammary epithelium, prostate  
 CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,  
 CC small airway epithelium, renal epithelium, umbilical artery smooth  
 CC muscle, neonatal dermal fibroblast, pulmonary artery smooth  
 CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,  
 CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,  
 CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,  
 CC osteoblasts or prostate stromal cell. The gene expression profile is used  
 CC for determining the level of RNA expression for a sample, determining the  
 CC phenotype of a cell and distinguishing cell types. The gene or a protein  
 CC expression profile is useful in identifying disease pathologies involving  
 CC alterations of gene expression. The assessment of expression profiles may  
 CC provide meaningful information with respect to tumour type and stage,  
 CC treatment methods, and prognosis. The gene or protein expression profile  
 CC may also be used for creating microarrays. The microarray is useful for  
 CC genetic and physical mapping of genomes. DNA sequencing, genetic or  
 CC medical diagnosis, genotyping of organisms, confirming cell or tissue  
 CC identifications and in identifying promising antibodies, antiviral or  
 CC antifungal agents  
 XX  
 SQ Sequence 4670 BP; 1072 A; 1094 C; 977 G; 1055 T; 0 U; 472 Other;  
 Query Match 100.0%; Score 20; DB 6; Length 4670;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGTCCATCTCATGTTGATG 20  
 DB 695 GGTCCATCTCATGTTGATG 714  
 RESULT 10  
 ABX05799  
 ID ABX05799 standard; DNA; 1029 BP.  
 XX  
 AC ABX05799;  
 XX  
 DT 27-OCT-2003 (revised)  
 XX  
 DT 11-FEB-2003 (first entry)  
 XX  
 DE S. pneumoniae type 4 strain coding region #87.  
 XX  
 KW Gene; ds; bacterial meningitis; pneumonia; sepsis; otitis media;  
 KW ear infection; antiinflammatory; antibacterial; immunostimulant;  
 KW auditory; respiratory; gene therapy; vaccine.  
 XX  
 OS Streptococcus pneumoniae; type 4 strain.  
 XX



WP ABS56454\_07 700001 810000  
WP ABS56454\_08 800001 910000  
WP ABS56454\_09 900001 1010000  
WP ABS56454\_10 1000001 1110000  
WP ABS56454\_11 1100001 1210000  
WP ABS56454\_12 1200001 1310000  
WP ABS56454\_13 1300001 1410000  
WP ABS56454\_14 1400001 1510000  
WP ABS56454\_15 1500001 1610000  
WP ABS56454\_16 1600001 1710000  
WP ABS56454\_17 1700001 1810000  
WP ABS56454\_18 1800001 1910000  
WP ABS56454\_19 1900001 2010000  
WP ABS56454\_20 2000001 2110000  
WP ABS56454\_21 2100001 2162598  
ID ABS56454 Standard; DNA; 2162598 BP.  
XX  
XX ABS56454;  
AC  
XX  
XX 27-OCT-2003 (revised)  
DT  
XX 10-FEB-2003 (first entry)  
DT  
XX  
XX Streptococcus pneumoniae type 4 strain complete genome.  
DE  
XX  
XX ds; bacterial meningitis; pneumonia; sepsis; otitis media; genome;  
KM ear infection; antiinflammatory; antibacterial; immunostimulant;  
KM auditory; respiratory; gene therapy; vaccine.  
XX  
XX OS Streptococcus pneumoniae; type 4 strain.  
XX  
XX WO200277021-A2.  
PN  
XX  
XX 03-OCT-2002.  
PD  
XX  
XX 27-MAR-2002; 2002WO-IB002163.  
PF  
XX  
XX 27-MAR-2001; 2001GB-00007658.  
PR  
XX  
XX (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Masignani V, Tettelin H, Fraser C;  
PI  
XX  
XX MPI; 2003-040579/03.  
DR  
XX  
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
PT useful as medicaments for treating or preventing a disease or infection  
PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or  
PT ear infection.  
XX  
XX  
XX Claim 17; SEQ ID NO 4979; 56bp; English.  
PS  
XX  
XX The invention relates to a protein comprising or having at least 50%  
CC identity to any of the 2469 amino acid sequences, identified in the  
CC specification (available on a computer readable format), or its fragment,  
CC expressed from 2469 of 2489 identified DNA coding regions from the  
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as  
CC ABS36454. Also included are an antibody which binds one of the proteins,  
CC treating a patient by administering the protein, DNA or antibody (in a  
CC composition), a kit comprising first and second primers, which are the  
CC nucleic acid cited above or fragments between nucleotides 8-100 of a  
CC sequence not defined in the specification, for amplifying a target  
CC sequence contained within a Streptococcus nucleic acid sequence, where  
CC the first primer is substantially complementary to the target sequence  
CC and the second primer is substantially complementary to the complement of  
CC the target sequence, and where the parts of the primers having  
CC substantial complementarity define the termini of the target sequence to  
CC be amplified, assay comprising contacting a test compound with the  
CC protein, and determining whether the test compound binds to the protein  
CC and a Streptococcus pneumoniae bacterium, where one or more genes  
CC encoding the proteins has been rendered inactive. The proteins, nucleic  
CC acid molecules, antibody and compositions are useful as medicaments for  
CC treating or preventing a disease or infection due to streptococcus

CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis  
CC media or ear infection. They are also useful in developing vaccines,  
CC diagnostics and antibiotics. The methods are useful for identifying  
CC immunodominant proteins. The present sequence is the Streptococcus  
CC pneumoniae type 4 strain genome sequence. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIP0 at  
CC ftp.wipo.int/pub/published\_pcr\_sequences. (Updated on 27-OCT-2003 to  
CC standardise OS field)  
XX  
XX  
SQ Sequence 2162598 BP; 654373A; 427176C; 431369G; 649680T; 0U; 00cher;  
Query Match 84.0%; Score 16.8; DB 7; Length 110000;  
Best Local Similarity 90.0%; Pred. No. 3.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GGTCATGCTCATGTGATG 20  
DB 67573 GGTCATGCTCATGTGATG 67554  
RESULT 13  
ID ABT13966/C  
XX ABT13966 standard; DNA; 6765 BP.  
XX  
AC ABT13966;  
XX  
XX 13-FEB-2003 (first entry)  
DT  
XX  
XX Human Cryptovirus strain BBR DNA region SEQ ID No 23.  
DE  
XX  
XX Neuroprotective; virucide; anticonvulsant; antiinflammatory; Cryptovirus;  
KM idiopathic; neurological; neurodegenerative; neuropsychological; vaccine;  
KM epilepsy; neuropsychiatric disorder; subacute sclerosing panencephalitis;  
KM epileptiform disease; multiple sclerosis; chronic fatigue syndrome;  
KM primary lymphadenopathy-associated illness; gene therapy; ds.  
XX  
XX OS Rubulavirus sp.  
XX  
XX WO200277211-A2.  
PN  
XX  
XX 03-OCT-2002.  
PD  
XX  
XX 07-FEB-2002; 2002WO-US004117.  
PF  
XX  
XX 07-FEB-2001; 2001US-0267253P.  
PR  
XX  
XX (CRYP-) CRYPTIC AFFILICTIONS LLC.  
PA  
XX  
XX Robbins SJ;  
PI  
XX  
XX MPI; 2003-040586/03.  
DR  
XX  
XX P-PSDB; ABO18512.  
DR  
XX  
XX New nucleic acid useful for diagnosing and treating idiopathic  
PT neurological disorders, including epileptiform diseases, e.g. epilepsy,  
PT and lymphadenopathy-associated illnesses, and in screening of potential  
PT new antiviral drugs.  
XX  
XX  
XX Disclosure; Page 225-235; 262pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising: contiguous  
CC nucleotide positions 1-15246 of a sequence comprising 15246 bp fully  
CC defined in the specification; a nucleotide sequence complementary to the  
CC 15246bp sequence; or a Cryptovirus-specific fragment of the 15246bp or  
CC its complement, at least about 5 nucleotides long. The nucleic acid is  
CC useful in diagnosing and treating many idiopathic neurological,  
CC neurodegenerative, neuropsychological and neuropsychiatric disorders,  
CC including epileptiform diseases (e.g. epilepsy, multiple sclerosis,  
CC chronic fatigue syndrome or subacute sclerosing panencephalitis) and  
CC primary lymphadenopathy-associated illnesses, and in research and  
CC development, including screening of potential new antiviral drugs. The  
CC nucleic acid, protein and the viral particle are useful in manufacturing

CC a vaccine. The protein is also used in producing a Cryptovirus-specific  
CC antibody. The antibody may also be used in manufacturing a medicament for  
CC the treatment of Cryptovirus infections. The polynucleotides of the  
CC invention can be used to treat disorders by gene therapy. This  
CC polynucleotide sequence represents a DNA region of the human Cryptovirus  
CC strain BBR of the invention

SQ Sequence 6765 BP; 2139 A; 1351 C; 1357 G; 1918 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 7; Length 6765;  
Best Local Similarity 94.4%; Pred. No. 3.3e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TCCATACCTCATGTTGATG 20  
Db 4712 TCCATACCTCATGTTGATG 4695

## RESULT 14

ABT13954/C  
ID ABT13954 standard; DNA; 15246 BP.

AC ABT13954;

DT 13-FEB-2003 (first entry)

DE Human Cryptovirus strain BBR coding sequence SEQ ID NO 1.

KW Neuroprotective; virucide; anticonvulsant; antiinflammatory; Cryptovirus;  
KW idiopathic; neurologic; neurodegenerative; neuropsychological; vaccine;  
KW epilepsy; neuropsychiatric disorder; subacute sclerosing panencephalitis;  
KW epileptiform disease; multiple sclerosis; chronic fatigue syndrome;  
KW primary lymphadenopathy-associated illness; gene therapy; gene; ds.

OS Rubulavirus sp.

PN WO20027211-A2.

PD 03-OCT-2002.

PF 07-FEB-2002; 2002WO-US004117.

PR 07-FEB-2001; 2001US-0267253P.

PA (CRYP-) CRYPTIC AFFLICTIONS LLC.

PI Robbins SJ;

DR WPI; 2003-040586/03.

PT New nucleic acid useful for diagnosing and treating idiopathic  
PT neurological disorders, including epileptiform diseases, e.g. epilepsy,  
PT and lymphadenopathy-associated illnesses, and in screening of potential  
PT new antiviral drugs.

PS Claim 1; Fig 9; 262pp; English.

XX The invention relates to an isolated nucleic acid comprising: contiguous  
CC nucleotide positions 1-15246 of a sequence comprising 15246 bp fully  
CC defined in the specification; a nucleotide sequence complementary to the  
CC 15246bp sequence; or a Cryptovirus-specific fragment of the 15246bp or  
CC its complement, at least about 5 nucleotides long. The nucleic acid is  
CC useful in diagnosing and treating many idiopathic neurological,  
CC neurodegenerative, neuropsychological and neuropsychiatric disorders,  
CC including epileptiform diseases (e.g. epilepsy, multiple sclerosis,  
CC chronic fatigue syndrome or subacute sclerosing panencephalitis) and  
CC primary lymphadenopathy-associated illnesses, and in research and  
CC development, including screening of potential new antiviral drugs. The  
CC nucleic acid, protein and the viral particle are useful in manufacturing  
CC a vaccine. The protein is also used in producing a Cryptovirus-specific  
CC antibody. The antibody may also be used in manufacturing a medicament for  
CC the treatment of Cryptovirus infections. The polynucleotides of the  
CC invention can be used to treat disorders by gene therapy. This

CC polynucleotide sequence represents the coding DNA for the human  
CC Cryptovirus strain BBR protein of the invention

SQ Sequence 15246 BP; 4753 A; 3273 C; 3131 G; 4089 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 7; Length 15246;  
Best Local Similarity 94.4%; Pred. No. 3.7e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TCCATACCTCATGTTGATG 20  
Db 13125 TCCATACCTCATGTTGATG 13108

## RESULT 15

ABA90521\_15  
Continuation (16 of 24) of ABA90521 from base 1500001 (Genomic sequence of Lactococcus 1  
WP Sequence split into 24 fragments LOCUS ABA90521 Accession ABA90521

WP Fragment Name Begin End

WP ABA90521\_00 1 110000

WP ABA90521\_01 100001 210000

WP ABA90521\_02 200001 310000

WP ABA90521\_03 300001 410000

WP ABA90521\_04 400001 510000

WP ABA90521\_05 500001 610000

WP ABA90521\_06 600001 710000

WP ABA90521\_07 700001 810000

WP ABA90521\_08 800001 910000

WP ABA90521\_09 900001 1010000

WP ABA90521\_10 1000001 1110000

WP ABA90521\_11 1100001 1210000

WP ABA90521\_12 1200001 1310000

WP ABA90521\_13 1300001 1410000

WP ABA90521\_14 1400001 1510000

WP ABA90521\_15 1500001 1610000

WP ABA90521\_16 1600001 1710000

WP ABA90521\_17 1700001 1810000

WP ABA90521\_18 1800001 1910000

WP ABA90521\_19 1900001 2010000

WP ABA90521\_20 2000001 2110000

WP ABA90521\_21 2100001 2210000

WP ABA90521\_22 2200001 2310000

WP ABA90521\_23 2300001 2365589

Query Match 82.0%; Score 16.4; DB 6; Length 110000;  
Best Local Similarity 94.4%; Pred. No. 4.8e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TCCATACCTCATGTTGATG 20  
Db 51686 TCCATACCTCATGTTGATG 51703

## RESULT 16

ABV97710/C  
ID ABV97710 standard; cDNA; 431 BP.

AC ABV97710;

DT 14-JAN-2003 (first entry)

DE Human pancreatic cancer expressed cDNA SEQ ID NO 3118.

KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;

KW cytostatic; tumour; gene; ss.

OS Homo sapiens.

PN WO200260317-A2.

PD 08-AUG-2002.

PF 30-JAN-2002; 2002WO-US002781.

```
XX 30-JAN-2001; 2001US-0265305P.
PR 31-JAN-2001; 2001US-0265305P.
PR 09-FEB-2001; 2001US-0267568P.
PR 21-MAR-2001; 2001US-0276551P.
PR 28-APR-2001; 2001US-0287112P.
PR 16-MAY-2001; 2001US-0291631P.
PR 12-JUL-2001; 2001US-0305484P.
PR 20-AUG-2001; 2001US-0313999P.
PR 27-NOV-2001; 2001US-0333626P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hapler WT, Jiang Y;
XX
XX WPI; 2002-627435/67.
DR
XX
XX New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT diagnosing, preventing and/or treating cancer, particularly pancreatic
PT cancer.
XX
XX Claim 1; SEQ ID NO 3118; 300bp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (1) comprising: (a)
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC complements of (a); (c) sequences consisting of at least 20 contiguous
CC residues of (a); (d) sequences that hybridize to (a), under moderately
CC stringent conditions; (e) sequences having at least 75% or 90% identity
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC ABP68633) encoded by (1) and oligonucleotide can be used to detect cancer
CC in a patient and compositions comprising polypeptides, polynucleotides,
CC antibodies, fusion proteins, T cell populations and antigen presenting
CC cells expressing the polypeptide are useful in treating pancreatic cancer
CC and stimulating an immune response. The polynucleotides can be used as
CC probes or primers for nucleic acid hybridisation, in the design and
CC preparation of ribozyme molecules for inhibiting expression of the tumour
CC polypeptides and proteins in the tumour cells, in vaccines and for gene
CC therapy. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published\_pct\_sequences
XX
SQ Sequence 431 BP; 150 A; 63 C; 60 G; 156 T; 0 U; 2 Other;

Query Match          79.0%; Score 15.8; DB 6; Length 431;
Best Local Similarity 89.5%; Pred. No. 4.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
   ||||| ||||| |||||
DB 194 GTCCATATATCAAGTGTATG 176

RESULT 17
ACH25004/C
ID ACH25004 standard; cDNA; 462 BP.
XX
XX ACH25004;
AC
XX 13-OCT-2003 (first entry)
DT
XX
DE Human adult ovary cDNA #3384.
XX
XX Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;
KM genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
OS
XX
XX US2003073623-A1.
PN
XX
XX 17-APR-2003.
PD
XX
XX 30-JUL-2001; 2001US-00918995.
PF
XX
```

```
PR 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
DR
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 12216; 44bp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SHH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=2003073623
XX
SQ Sequence 462 BP; 104 A; 120 C; 124 G; 112 T; 0 U; 2 Other;

Query Match          79.0%; Score 15.8; DB 8; Length 462;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20
   ||||| ||||| |||||
DB 446 GTCCCTTCATGTTGATG 428

RESULT 18
AAA16057
ID AAA16057 standard; DNA; 614 BP.
XX
XX AAA16057;
AC
XX 14-JUN-2000 (first entry)
DT
XX
DE Human colon cancer differentially expressed nucleotide sequence #62.
XX
XX Colon cancer; detect; differential expression; human; treatment;
KM detect mutation; non-invasive diagnostic method; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200012702-A2.
PN
XX
XX 09-MAR-2000.
PD
XX
XX 30-AUG-1999; 99WO-US019424.
PF
XX
XX 31-AUG-1998; 98US-0098639P.
PR 27-JAN-1999; 99US-0117393P.
XX
XX (FARB ) BAYER CORP.
PA
```



XX Endege WO, Steinmann KE, Astle JH, Burgess CC, Carroll E;  
 PI Carino TJ, Divedi P, Ford DM, Lewis ME, Molino GA, Monahan JB;  
 PI Schlegel R;  
 DR WPI; 2000-256641/22.  
 XX Novel nucleic acids and proteins for identifying therapeutic agents  
 PT useful for treating and diagnosing cancer, especially colon cancer.  
 PS Claim 16; Page 149; 345pp; English.  
 XX This sequence represents a human nucleotide sequence which is  
 CC differentially expressed in colon cancer cells compared to the expression  
 CC levels in normal cells. The nucleotide sequence can be used as a source  
 CC of primers and probes. The nucleotide sequence is useful for determining  
 CC the phenotype of a cell by detecting the differential expression of the  
 CC sequence relative to a normal cell. The probes derived from the sequence  
 CC can also be used to determine the phenotype of cells in a sample. Probes  
 CC and antibodies which hybridise to the nucleotide sequence can also be  
 CC used to determine the phenotype of a cell. The primers are useful for  
 CC detecting a mutation in a test nucleotide sequence and also for detecting  
 CC cancer, preferably colon cancer. Antibodies against the protein encoded  
 CC by the nucleotide sequence can also be used in a method to detect colon  
 CC cancer. The diagnostic method is non-invasive and accurate for diagnosing  
 CC colon cancer at an early stage  
 SQ Sequence 614 BP; 202 A; 91 C; 99 G; 214 T; 0 U; 8 Other;  
 OY Query Match 79.0%; Score 15.8; DB 3; Length 614;  
 Best Local Similarity 89.5%; Pred. No. 4.8e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 DB 2 GTCCATCTCATGTTGATG 20  
 466 GTCCATATCAAGTGATG 484  
 RESULT 19  
 ABL23928  
 ID ABL23928 standard; DNA; 2830 BP.  
 XX ABL23928;  
 AC  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 23257.  
 DE  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; de.  
 KM  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX Claim 1; SEQ ID NO 23257; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-  
 CC AB872072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 SQ Sequence 2830 BP; 740 A; 666 C; 674 G; 750 T; 0 U; 0 Other;  
 OY Query Match 79.0%; Score 15.8; DB 4; Length 2830;  
 Best Local Similarity 89.5%; Pred. No. 6e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 DB 1 GATCCATCTCATGTTGAT 19  
 2480 GGTTCATAGTCATGTTGAT 2498  
 RESULT 20  
 ABL11966/c  
 ID ABL11966 standard; cDNA; 6454 BP.  
 XX ABL11966;  
 AC  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 30380.  
 DE  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 KM  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR P-PsDB; ABB67863.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX Claim 1; SEQ ID NO 30380; 21pp + Sequence Listing; English.  
 PS  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA  
 CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-  
 CC AB872072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 SQ Sequence 6454 BP; 1917 A; 1358 C; 1277 G; 1902 T; 0 U; 0 Other;

Query Match 79.0%; Score 15.8; DB 4; Length 6454;  
 Best Local Similarity 89.5%; Pred. No. 6.7e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GTCCTACTCATGTTGATG 20  
 |||||  
 DB 1006 GTTCATAGTCATGATGATG 988

## RESULT 21

ABL03787/c  
 ID ABL03787 standard; cDNA; 8114 BP.

AC ABL03787;  
 XX

DT 26-MAR-2002 (first entry)  
 XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5843.  
 XX

KM Drosophila; developmental biology; cell signalling; insecticide;  
 XX

KW pharmaceutical; gene; ss.  
 XX

OS Drosophila melanogaster.  
 XX

PN WO200171042-A2.  
 XX

PD 27-SEP-2001.  
 XX

PF 23-MAR-2001; 2001WO-US009231.  
 XX

PR 23-MAR-2000; 2000US-0191637P.  
 XX

PP 11-JUL-2000; 2000US-00614150.  
 XX

PA (PEKE ) PE CORP NY.  
 XX

PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX

DR WPI; 2001-656860/75.  
 XX

P-PsDB; ABB59684.  
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX genes from Drosophila and for elucidating cell signalling and cell-cell  
 XX interactions.  
 XX

PS Claim 1; SEQ ID NO 5843; 21pp + Sequence Listing; English.  
 XX

CC The invention relates to an isolated nucleic acid detection reagent  
 XX capable of detecting 1000 or more genes from Drosophila. The invention is  
 XX useful in developmental biology and in elucidating cell signalling and  
 XX cell-cell interactions in higher eukaryotes for the development of  
 XX insecticides, therapeutics and pharmaceutical drugs. The invention  
 XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 XX ABB72072). The sequence data for this patent did not form part of the  
 XX printed specification, but was obtained in electronic format directly  
 XX from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences  
 XX

SO Sequence 8114 BP; 2159 A; 2177 C; 2263 G; 1515 T; 0 U; 0 Other;  
 XX

Query Match 79.0%; Score 15.8; DB 4; Length 8114;  
 Best Local Similarity 89.5%; Pred. No. 6.9e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGTCCATCTCATGTTGAT 19  
 |||||  
 DB 7755 GGTCATAGTCATGTTGAT 7737

## RESULT 22

ABL03786  
 ID ABL03786 standard; cDNA; 18737 BP.  
 XX

AC ABL03786;  
 XX

XX 26-MAR-2002 (first entry)  
 DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 5840.  
 XX

DE Drosophila; developmental biology; cell signalling; insecticide;  
 XX

KM Drosophila; developmental biology; cell signalling; insecticide;  
 XX

KW pharmaceutical; gene; ss.  
 XX

OS Drosophila melanogaster.  
 XX

PN WO200171042-A2.  
 XX

PD 27-SEP-2001.  
 XX

PF 23-MAR-2001; 2001WO-US009231.  
 XX

PR 23-MAR-2000; 2000US-0191637P.  
 XX

PP 11-JUL-2000; 2000US-00614150.  
 XX

PA (PEKE ) PE CORP NY.  
 XX

PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX

DR WPI; 2001-656860/75.  
 XX

P-PsDB; ABB59683.  
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX genes from Drosophila and for elucidating cell signalling and cell-cell  
 XX interactions.  
 XX

PS Claim 1; SEQ ID NO 5840; 21pp + Sequence Listing; English.  
 XX

CC The invention relates to an isolated nucleic acid detection reagent  
 XX capable of detecting 1000 or more genes from Drosophila. The invention is  
 XX useful in developmental biology and in elucidating cell signalling and  
 XX cell-cell interactions in higher eukaryotes for the development of  
 XX insecticides, therapeutics and pharmaceutical drugs. The invention  
 XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 XX ABB72072). The sequence data for this patent did not form part of the  
 XX printed specification, but was obtained in electronic format directly  
 XX from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences  
 XX

SO Sequence 18737 BP; 4552 A; 4551 C; 4677 G; 4957 T; 0 U; 0 Other;  
 XX

Query Match 79.0%; Score 15.8; DB 4; Length 18737;  
 Best Local Similarity 89.5%; Pred. No. 7.7e+02;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GGTCCATCTCATGTTGAT 19  
 |||||  
 DB 1360 GGTCATAGTCATGTTGAT 1378

## RESULT 23

ABL0396/c  
 ID ABL0396 standard; cDNA; 24401 BP.  
 XX

AC ABL0396;  
 XX

DT 26-MAR-2002 (first entry)  
 XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 4670.  
 XX

KM Drosophila; developmental biology; cell signalling; insecticide;  
 XX

KW pharmaceutical; gene; ss.  
 XX

OS Drosophila melanogaster.  
 XX

PN WO200171042-A2.  
 XX

PD 27-SEP-2001.  
 XX

PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PMD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
DR P-PSDB; ABB59293.  
XX  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
XX  
XX Claim 1; SEQ ID NO 4670; 21bp + Sequence Listing; English.  
PS  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 24401 BP; 7446 A; 4747 C; 4751 G; 7457 T; 0 U; 0 Other;  
XX  
XX  
XX Query Match 79.0%; Score 15.8; DB 4; Length 24401;  
XX Best Local Similarity 89.5%; Pred. No. 8e+02;  
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
Oy 2 GTCCATCTCATGTGATG 20  
Db 12878 GTCAATCTCATGTGATG 12860  
XX  
XX  
XX RESULT 24  
XX ABK42099  
XX ID ABK42099 standard; CDNA; 551 BP.  
XX AC ABK42099;  
XX  
XX DT 21-MAY-2002 (first entry)  
XX  
XX DE CDNA encoding novel human connective tissue related polypeptide #487.  
XX  
XX KW Human; connective tissue related disorder; cancer; gene therapy;  
XX cytosolic; gene; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200155343-A1.  
XX  
XX PD 02-AUG-2001.  
XX  
XX PF 17-JAN-2001; 2001WO-US001322.  
XX  
XX PR 31-JAN-2000; 2000US-0179065P.  
XX PR 04-FEB-2000; 2000US-0180628P.  
XX PR 24-FEB-2000; 2000US-0184664P.  
XX PR 02-MAR-2000; 2000US-0186350P.  
XX PR 16-MAR-2000; 2000US-0189874P.  
XX PR 17-MAR-2000; 2000US-0190076P.  
XX PR 18-APR-2000; 2000US-0198123P.  
XX PR 19-MAY-2000; 2000US-020515P.  
XX PR 07-JUN-2000; 2000US-0209467P.  
XX PR 28-JUN-2000; 2000US-0214886P.  
XX PR 30-JUN-2000; 2000US-0215135P.  
XX PR 07-JUL-2000; 2000US-0216647P.  
XX

PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225279P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0232887P.  
PR 01-SEP-2000; 2000US-0232934P.  
PR 01-SEP-2000; 2000US-0232934P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
XX

PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX  
DR WPI; 2001-565190/63.  
XX  
XX  
PT P-PSDB; AAU86921.  
XX  
XX  
PT Nucleic acid encoding novel connective tissue associated polypeptides,  
XX  
XX  
PT used in diagnosing, preventing, treating or ameliorating a disorder such  
XX  
XX  
PT as cancer or rheumatoid arthritis.  
XX  
XX  
PS Claim 4; SEQ ID NO 497; 673bp; English.  
XX  
XX  
XX The present invention relates to the isolation of novel human connective  
XX  
XX tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide  
XX  
XX (cDNA and genomic) sequences encoding them. The sequences of the  
XX  
XX invention are useful in the diagnosis, treatment, prevention and/or  
XX  
XX prognosis of diseases associated with connective tissue(s), including  
XX  
XX cancer. The polynucleotide sequences of the invention are also useful in  
XX  
XX gene therapy. ABK41613-ABK42101 represent cDNA sequences encoding the  
XX  
XX novel human connective tissue related polypeptides. Note: The sequence

CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 551 BP; 143 A; 114 G; 194 T; 0 U; 6 Other;  
Query Match 77.0%; Score 15.4; DB 4; Length 551;  
Best Local Similarity 94.1%; Pred. No. 7.66+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 3 TCACATCATGTTGAT 19  
|||||  
Db 261 TCATCTCATGTTGAT 277  
|||||  
RESULT 25  
ADB59766 standard; cDNA, 551 BP.  
XX  
XX ADB59766;  
XX  
AC 04-DEC-2003 (first entry)  
XX  
DT  
XX  
DE Connective tissue related polynucleotide #487.  
XX  
XX cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;  
XX antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;  
XX antiinflammatory; anti-allergic; antiasthmatic; dermatological;  
XX nephropoc; virucide; fungicide; antibacterial; antiparasitic;  
XX gene therapy; ds; connective tissues disorder; rheumatoid arthritis;  
XX systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;  
XX cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;  
XX Alzheimer's disease; Parkinson's disease; cardiovascular disease;  
XX atherosclerosis; myocarditis; cardiopulmonary bypass complication;  
XX autoimmune disease; multiple sclerosis; allergic reaction; asthma;  
XX rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;  
XX gastrointestinal disorder; inflammatory bowel disease;  
XX organ transplant rejection; immune system disorder; Bruton's disease;  
XX X-linked lymphoproliferative syndrome;  
XX B-cell lymphoproliferative disorder; HIV; AIDS; infection;  
XX chromosome identification; chromosome mapping;  
XX connective tissue related polynucleotide; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX OS  
XX  
XX PN US2003054375-A1.  
XX  
XX PD 20-MAR-2003.  
XX  
XX  
XX 07-MAR-2002; 2002US-00092154.  
XX  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.

PR	14-AUG-2000	2000US-02252666
PR	14-AUG-2000	2000US-02252676
PR	14-AUG-2000	2000US-02252686
PR	14-AUG-2000	2000US-02252707
PR	14-AUG-2000	2000US-02252737
PR	14-AUG-2000	2000US-02252747
PR	14-AUG-2000	2000US-02252757
PR	14-AUG-2000	2000US-02252768
PR	14-AUG-2000	2000US-02252788
PR	14-AUG-2000	2000US-02252799
PR	18-AUG-2000	2000US-02262719
PR	18-AUG-2000	2000US-02262681
PR	22-AUG-2000	2000US-02266688
PR	22-AUG-2000	2000US-02271828
PR	23-AUG-2000	2000US-02270099
PR	30-AUG-2000	2000US-02289249
PR	01-SEP-2000	2000US-02292878
PR	01-SEP-2000	2000US-02293439
PR	01-SEP-2000	2000US-02293449
PR	01-SEP-2000	2000US-02293459
PR	05-SEP-2000	2000US-02295539
PR	05-SEP-2000	2000US-02295139
PR	06-SEP-2000	2000US-02304379
PR	06-SEP-2000	2000US-02304389
PR	08-SEP-2000	2000US-02312429
PR	08-SEP-2000	2000US-02312439
PR	08-SEP-2000	2000US-02312449
PR	08-SEP-2000	2000US-02314139
PR	08-SEP-2000	2000US-02314149
PR	08-SEP-2000	2000US-02314159
PR	08-SEP-2000	2000US-02320809
PR	12-SEP-2000	2000US-02319688
PR	14-SEP-2000	2000US-02323379
PR	14-SEP-2000	2000US-02323389
PR	14-SEP-2000	2000US-02323399
PR	14-SEP-2000	2000US-02324009
PR	14-SEP-2000	2000US-02324019
PR	14-SEP-2000	2000US-02330639
PR	14-SEP-2000	2000US-02330649
PR	14-SEP-2000	2000US-02330659
PR	21-SEP-2000	2000US-02342239
PR	21-SEP-2000	2000US-02342749
PR	21-SEP-2000	2000US-02342759
PR	25-SEP-2000	2000US-02349979
PR	25-SEP-2000	2000US-02349989
PR	26-SEP-2000	2000US-02355849
PR	26-SEP-2000	2000US-02355834
PR	27-SEP-2000	2000US-02355839
PR	27-SEP-2000	2000US-02355869
PR	29-SEP-2000	2000US-02363679
PR	29-SEP-2000	2000US-02363689
PR	29-SEP-2000	2000US-02363699
PR	29-SEP-2000	2000US-02363709
PR	02-OCT-2000	2000US-02368029
PR	02-OCT-2000	2000US-02370379
PR	02-OCT-2000	2000US-02370389
PR	02-OCT-2000	2000US-02370399
PR	13-OCT-2000	2000US-02393959
PR	13-OCT-2000	2000US-02393979
PR	20-OCT-2000	2000US-02405609
PR	20-OCT-2000	2000US-02412219
PR	20-OCT-2000	2000US-02411859
PR	20-OCT-2000	2000US-02411869
PR	20-OCT-2000	2000US-02417879
PR	20-OCT-2000	2000US-02418089
PR	20-OCT-2000	2000US-02418099
PR	01-NOV-2000	2000US-02418269
PR	01-NOV-2000	2000US-02446179
PR	08-NOV-2000	2000US-02464749
PR	08-NOV-2000	2000US-02464759
PR	08-NOV-2000	2000US-02464769
PR	08-NOV-2000	2000US-02464779
PR	08-NOV-2000	2000US-02464789
PR	08-NOV-2000	2000US-02465239
PR	08-NOV-2000	2000US-02465249

XX	PR	08-NOV-2000	2000US-0246525P	
XX	PR	08-NOV-2000	2000US-0246526P	
XX	PR	08-NOV-2000	2000US-0246527P	
XX	PR	08-NOV-2000	2000US-0246528P	
XX	PR	08-NOV-2000	2000US-0246532P	
XX	PR	08-NOV-2000	2000US-0246538P	
XX	PR	08-NOV-2000	2000US-0246609P	
XX	PR	08-NOV-2000	2000US-0246610P	
XX	PR	08-NOV-2000	2000US-0246611P	
XX	PR	08-NOV-2000	2000US-0246613P	
XX	PR	17-NOV-2000	2000US-0249207P	
XX	PR	17-NOV-2000	2000US-0249208P	
XX	PR	17-NOV-2000	2000US-0249209P	
XX	PR	17-NOV-2000	2000US-0249210P	
XX	PR	17-NOV-2000	2000US-0249211P	
XX	PR	17-NOV-2000	2000US-0249212P	
XX	PR	17-NOV-2000	2000US-0249213P	
XX	PR	17-NOV-2000	2000US-0249214P	
XX	PR	17-NOV-2000	2000US-0249215P	
XX	PR	17-NOV-2000	2000US-0249216P	
XX	PR	17-NOV-2000	2000US-0249217P	
XX	PR	17-NOV-2000	2000US-0249218P	
XX	PR	17-NOV-2000	2000US-0249244P	
XX	PR	17-NOV-2000	2000US-0249245P	
XX	PR	17-NOV-2000	2000US-0249246P	
XX	PR	17-NOV-2000	2000US-0249265P	
XX	PR	17-NOV-2000	2000US-0249265P	
XX	PR	17-NOV-2000	2000US-0249297P	
XX	PR	17-NOV-2000	2000US-0249299P	
XX	PR	17-NOV-2000	2000US-0249300P	
XX	PR	01-DEC-2000	2000US-0250160P	
XX	PR	01-DEC-2000	2000US-0250391P	
XX	PR	05-DEC-2000	2000US-0251030P	
XX	PR	05-DEC-2000	2000US-0251988P	
XX	PR	05-DEC-2000	2000US-0256719P	
XX	PR	06-DEC-2000	2000US-0251479P	
XX	PR	08-DEC-2000	2000US-0251856P	
XX	PR	08-DEC-2000	2000US-0251866P	
XX	PR	08-DEC-2000	2000US-0251869P	
XX	PR	08-DEC-2000	2000US-0251989P	
XX	PR	08-DEC-2000	2000US-0251990P	
XX	PR	11-DEC-2000	2000US-0254097P	
XX	PR	05-JAN-2001	2001US-0259678P	
XX	PR	17-JAN-2001	2001US-00764847	
XX	PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX	PI			
XX	PI	Rosen CA,	Ruben SM, Barash SC;	
XX	PI	WPI: 2003-634869/60.		
XX	DR	P-PSDB; ADB60255.		
XX	PT			
XX	PT	New connective tissue-related polypeptides and polynucleotides, useful		
XX	PT	for treating, preventing and/or prognosing e.g. disorders of connective		
XX	PT	tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or		
XX	PT	neoplasias.		
XX	PS	Claim 1; SEQ ID NO 497; 248pp. English.		
XX	XX			
XX	XX	The invention describes an isolated nucleic acid molecule (1), which		
XX	CC	comprises a sequence that is at least 95 % identical to a connective		
XX	CC	tissue-related polynucleotide encoding connective tissue antigens (CTA)		
XX	CC	The polypeptide or polynucleotide is useful for preventing, treating, or		
XX	CC	ameliorating medical conditions in a mammal. The connective tissue		
XX	CC	polypeptides, polynucleotides and antibodies are particularly useful for		
XX	CC	treating, preventing and/or prognosing disorders of connective tissues		
XX	CC	(e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,		
XX	CC	scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or		
XX	CC	neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.		
XX	CC	Alzheimer's disease, or Parkinson's disease), cardiovascular diseases		
XX	CC	(e.g. atherosclerosis, myocarditis or cardiopulmonary bypass		
XX	CC	complications), autoimmune diseases (e.g. systemic lupus erythematosus,		
XX	CC	rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.		

Best Local Similarity 94.1%; Pred. No. 7.6e+02;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGAT 19  
| | | | | | | | | | | | | | | | | | | | | |  
DB 261 TCCATTCATCATGTTGAT 277

## RESULT 26

ABV55825/c  
ID ABV55825 standard; cDNA, 609 BP.

XX ABV55825;  
AC

XX 17-SEP-2002 (first entry)  
DT

XX Human prostate expression marker cDNA 55816.  
DE

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW

XX pharmacogenomic marker; gene; ss.  
XX

OS Homo sapiens.  
XX

XX WO200160860-A2.  
PN

XX 23-AUG-2001.  
PD

XX 20-FEB-2001; 2001WO-US005171.  
PF

XX 17-FEB-2000; 2000US-018319P.  
PR

XX 16-MAR-2000; 2000US-0189862P.  
PR

XX 25-MAY-2000; 2000US-0207454P.  
PR

XX 09-JUN-2000; 2000US-0211314P.  
PR

XX 18-JUL-2000; 2000US-0219007P.  
PR

XX 13-DEC-2000; 2000US-0255281P.  
PR

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA

XX Schlegel R, Endege WO, Monahan JR;  
PI

XX WPI, 2001-662795/76.  
DR

XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 10778; 11/50pp; English.  
PS

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or incidence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX

XX Sequence 609 BP; 211 A; 104 C; 87 G; 204 T; 0 U; 3 Other;  
SQ

Query Match 77.0%; Score 15.4; DB 5; Length 609;  
Best Local Similarity 94.1%; Pred. No. 7.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCCATACATCATGTTGAT 19  
| | | | | | | | | | | | | | | | | | | | | |  
DB 468 TCCATACATCATGTTGAT 452

## RESULT 27

AAH52789  
ID AAH52789 standard; DNA; 633 BP.  
XX

XX AAH52789;  
AC

XX 03-SEP-2001 (first entry)  
DT

XX S. epidermidis open reading frame nucleotide sequence SEQ ID NO:971.  
DE

XX Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
KW

XX endocarditis; ds.  
KW

XX Staphylococcus epidermidis.  
OS

XX WO200134809-A2.  
PN

XX 17-MAY-2001.  
PD

XX 09-NOV-2000; 2000WO-US030782.  
PF

XX 09-NOV-1999; 99US-0164258P.  
PR

XX (GLAXO ) GLAXO GROUP LTD.  
PA

XX Kimmeryly WJ;  
PI

XX WPI, 2001-316495/33.  
DR

XX P-PsDB; AAG81939.  
DR

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
PT useful for vaccinating against infections, e.g. endocarditis.

XX Claim 8; Page 285; 218pp; English.  
PS

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)  
CC and (II) can have antibacterial activity and therefore can be used in  
CC vaccination. The nucleic acids (I) may be used to produce the S.  
CC epidermidis polypeptides (II) via the production of vectors containing  
CC them which are used to produce host cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the  
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to  
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA  
CC polynucleotide sequences from the present invention. AAH5091 to AAH5098  
CC represent oligonucleotide sequences and primers which are used in the  
CC exemplification of the present invention. N.B. The present invention  
CC specifically claims all the polynucleotide sequences given in the  
CC sequence listing of the present specification, however the sequence  
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present  
CC for SEQ ID NO:4455 to 4464  
XX

XX Sequence 633 BP; 238 A; 87 C; 124 G; 184 T; 0 U; 0 Other;  
SQ

Query Match 77.0%; Score 15.4; DB 4; Length 633;  
Best Local Similarity 94.1%; Pred. No. 7.7e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCACTCATGTTGA 18  
| | | | | | | | | | | | | | | | | | | | | |  
DB 385 GTCCATCACTCATGTTGA 401

## RESULT 28

AAH53645  
ID AAH53645 standard; DNA; 675 BP.  
XX

XX AAH53645;  
AC

XX 03-SEP-2001 (first entry)  
DT

XX	S.	epidermidis open reading frame nucleotide sequence SEQ ID NO:2683.
XX	Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;	
KW	endocarditis; ds.	
XX		
OS	Staphylococcus epidermidis.	
PN	WO200134809-A2.	
PD	17-MAY-2001.	
PR	09-NOV-2000; 2000MO-US030782.	
PA	09-NOV-1999; 99US-0164258P.	
PB	(GLAX ) GLAXO GROUP LTD.	
PI	Kimmerly WJ;	
DR	WPI, 2001-316495/33.	
PT	P-PsDB; AAG82795.	
PS	Nucleic acids encoding polypeptides from Staphylococcus epidermidis,	
CC	useful for vaccinating against infections, e.g. endocarditis.	
CC	Claim 8; Page 702; 218pp; English.	
CC	AH52304 to AH53970 represent nucleic acids (I) encoding polypeptides	
CC	(II), given in AAG8154 to AAG83120, from Staphylococcus epidermidis. (I)	
CC	and (II) can have antibacterial activity and therefore can be used in	
CC	vaccination. The nucleic acids (I) may be used to produce the S.	
CC	epidermidis polypeptides (II) via the production of vectors containing	
CC	them which are used to produce hosts cells which express the	
CC	polypeptides. The polypeptides (II) (and/or nucleic acids) may then be	
CC	used to vaccinate subjects and to raise antibodies against the bacteria.	
CC	The polypeptides may also be used to assay for other inhibitors of their	
CC	activity and therefore identify compounds that may be used for the	
CC	treatment of S. epidermidis infections, e.g. endocarditis. AH53971 to	
CC	AH55090 represent specifically claimed S. epidermidis genomic DNA	
CC	polynucleotide sequences from the present invention. AH55091 to AH55098	
CC	represent oligonucleotide sequences and primers which are used in the	
CC	exemplification of the present invention. N.B. The present invention	
CC	specifically claims all the polynucleotide sequences given in the	
CC	sequence listing of the present specification, however the sequence	
CC	listing only goes up to SEQ ID NO:4454 so even though sequences are given	
CC	in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present	
CC	for SEQ ID NO:4455 to 4464	
CC		
XX		
SQ	Sequence 675 BP; 260 A; 91 C; 132 G; 192 T; 0 U; 0 Other;	
	Query Match	77.0%; Score 15.4; DB 4; Length 675;
	Best Local Similarity	94.1%; Pred. No. 7.8e+02;
	Matches 16; Conservative	0; Mismatches 1; Indels 0; Gaps 0
OY	2 GTGCATACGATGTGA 18	
Db	427 GTCGATTCTCATGTTGA 443	
RESULT 29		
ID	ABN90925	
XX	ABN90925 standard; DNA; 675 BP.	
XX		
AC	ABN90925;	
XX		
DT	24-JUL-2002 (first entry)	
DE	Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:388.	
KW	Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;	
XX	antibacterial; gene therapy; gene; ds.	
XX		

OS	Staphylococcus epidermidis.
XX	
XX	US6380370-B1.
XX	
PD	30-APR-2002.
XX	
PF	13-AUG-1998; 98US-00134001.
XX	
XX	14-AUG-1997; 97US-0055779P.
PR	08-NOV-1997; 97US-0064964P.
XX	
PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	
PI	Doucette-Stamm LA, Bush D;
DR	WPI; 2002-381255/41.
DR	P-PSDB; ABP38380.
XX	
PT	Novel isolated nucleic acid encoding a Staphylococcus epidermis
PT	polypeptide, useful for diagnosing and treating bacterial infections.
XX	
PS	Disclosure; SEQ ID NO 388; 267pp; English.
CC	
CC	ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC	frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC	given in ABP5124 to ABP37960. The S. epidermidis sequences have
CC	antibacterial activity and can be used in gene therapy. The sequences can
CC	also be used in the diagnosis and treatment of bacterial infections,
CC	particularly S. epidermidis infections. The sequences can be used to
CC	screen for compounds able to interfere with the S. epidermidis life cycle
CC	or inhibit S. epidermidis infection. N.B. The sequence data for this
CC	patent did not form part of the printed specification, but was obtained
CC	in electronic format directly from the USPro web site
XX	
SQ	Sequence 675 BP; 260 A; 90 C; 132 G; 193 T; 0 U; 0 Other;
	Query Match 77.0%; Score 15.4; DB 6; Length 675;
	Best Local Similarity 94.1%; Pred.No. 7.8e+02;
	Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	2 GTCCATCTCATGTTGA 18 
DB	427 GTCGATCTCATGTTGA 443
RESULT 30	
AAC46198/c	
ID	AAC46198 standard; DNA; 2132 BP.
XX	
AC	AAC46198;
XX	
DT	18-OCT-2000 (first entry)
XX	
EE	Arabidopsis thaliana DNA fragment SEQ ID NO: 49255.
XX	
KW	Hybridisation assay; generic mapping; gene expression control;
KW	protein identification; signal transduction pathway; metabolic pathway;
XX	promoter; termination sequence; ss.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439.
XX	
PR	25-FEB-1999; 99US-0121825P.
PR	05-MAR-1999; 99US-0123180P.
PR	09-MAR-1999; 99US-0123548P.
PR	23-MAR-1999; 99US-0125788P.
PR	25-MAR-1999; 99US-0126264P.
PR	29-MAR-1999; 99US-0126785P.

PR 01-APR-1999;	99US-0127462P.	PR 19-JUL-1999;	99US-0144331P.
PR 06-APR-1999;	99US-0128234P.	PR 19-JUL-1999;	99US-0144332P.
PR 08-APR-1999;	99US-0128714P.	PR 19-JUL-1999;	99US-0144333P.
PR 16-APR-1999;	99US-0129845P.	PR 19-JUL-1999;	99US-0144334P.
PR 19-APR-1999;	99US-0130077P.	PR 19-JUL-1999;	99US-0144335P.
PR 21-APR-1999;	99US-0130449P.	PR 20-JUL-1999;	99US-0144335P.
PR 23-APR-1999;	99US-0130510P.	PR 20-JUL-1999;	99US-0144632P.
PR 23-APR-1999;	99US-0130891P.	PR 20-JUL-1999;	99US-0144884P.
PR 28-APR-1999;	99US-0131449P.	PR 21-JUL-1999;	99US-0144814P.
PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145066P.
PR 30-APR-1999;	99US-0132407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-0132484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-0132485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-0132486P.	PR 22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999;	99US-0132487P.	PR 22-JUL-1999;	99US-0145192P.
PR 07-MAY-1999;	99US-0132863P.	PR 23-JUL-1999;	99US-0145145P.
PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134368P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135629P.	PR 03-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141827P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142977P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144086P.	PR 07-OCT-1999;	99US-0158029P.
PR 19-JUL-1999;	99US-0144325P.	PR 08-OCT-1999;	99US-0158232P.
		PR 12-OCT-1999;	99US-0158369P.





PR	21-JUL-1999	99US-0144884P
PR	21-JUL-1999	99US-0144814P
PR	21-JUL-1999	99US-0145086P
PR	21-JUL-1999	99US-0145088P
PR	22-JUL-1999	99US-0145085P
PR	22-JUL-1999	99US-0145087P
PR	22-JUL-1999	99US-0145089P
PR	22-JUL-1999	99US-0145192P
PR	23-JUL-1999	99US-0145185P
PR	23-JUL-1999	99US-0145128P
PR	26-JUL-1999	99US-0145224P
PR	26-JUL-1999	99US-0145276P
PR	27-JUL-1999	99US-0145513P
PR	27-JUL-1999	99US-0145518P
PR	27-JUL-1999	99US-0145519P
PR	28-JUL-1999	99US-0145591P
PR	28-JUL-1999	99US-0145638P
PR	02-AUG-1999	99US-0146388P
PR	02-AUG-1999	99US-0146389P
PR	03-AUG-1999	99US-0147038P
PR	04-AUG-1999	99US-0147204P
PR	04-AUG-1999	99US-0147302P
PR	05-AUG-1999	99US-0147315P
PR	05-AUG-1999	99US-0147762P
PR	06-AUG-1999	99US-0147703P
PR	06-AUG-1999	99US-0147703P
PR	09-AUG-1999	99US-0147943P
PR	09-AUG-1999	99US-0147946P
PR	10-AUG-1999	99US-0148173P
PR	10-AUG-1999	99US-0148319P
PR	11-AUG-1999	99US-0148319P
PR	12-AUG-1999	99US-0148341P
PR	13-AUG-1999	99US-0148565P
PR	13-AUG-1999	99US-0148684P
PR	16-AUG-1999	99US-0149368P
PR	17-AUG-1999	99US-01494175P
PR	18-AUG-1999	99US-0149426P
PR	20-AUG-1999	99US-0149722P
PR	20-AUG-1999	99US-0149723P
PR	20-AUG-1999	99US-0149829P
PR	23-AUG-1999	99US-0149930P
PR	23-AUG-1999	99US-0149930P
PR	25-AUG-1999	99US-0150566P
PR	26-AUG-1999	99US-0150864P
PR	27-AUG-1999	99US-0151065P
PR	27-AUG-1999	99US-0151066P
PR	27-AUG-1999	99US-0151080P
PR	30-AUG-1999	99US-0151103P
PR	31-AUG-1999	99US-0151438P
PR	01-SEP-1999	99US-0151930P
PR	07-SEP-1999	99US-0152363P
PR	10-SEP-1999	99US-0153070P
PR	13-SEP-1999	99US-0153758P
PR	15-SEP-1999	99US-0154018P
PR	16-SEP-1999	99US-0154039P
PR	20-SEP-1999	99US-0154719P
PR	22-SEP-1999	99US-0155131P
PR	23-SEP-1999	99US-0155466P
PR	24-SEP-1999	99US-0155659P
PR	28-SEP-1999	99US-0156488P
PR	29-SEP-1999	99US-0156566P
PR	04-OCT-1999	99US-0157117P
PR	05-OCT-1999	99US-0157753P
PR	06-OCT-1999	99US-0157665P
PR	07-OCT-1999	99US-0158029P
PR	08-OCT-1999	99US-0158232P
PR	12-OCT-1999	99US-0158369P
PR	13-OCT-1999	99US-0159293P
PR	13-OCT-1999	99US-0159294P
PR	13-OCT-1999	99US-0159295P
PR	14-OCT-1999	99US-0159329P
PR	14-OCT-1999	99US-0159330P
PR	14-OCT-1999	99US-0159331P
PR	14-OCT-1999	99US-0159637P

PR	14-Oct-1939	99US-0159658P
PR	19-Oct-1939	99US-0159654P
PR	21-Oct-1939	99US-0160741P
PR	21-Oct-1939	99US-0160767P
PR	21-Oct-1939	99US-0160768P
PR	21-Oct-1939	99US-0160770P
PR	21-Oct-1939	99US-0160814P
PR	21-Oct-1939	99US-0160815P
PR	22-Oct-1939	99US-0160980P
PR	22-Oct-1939	99US-0160981P
PR	22-Oct-1939	99US-0160982P
PR	25-Oct-1939	99US-0161404P
PR	25-Oct-1939	99US-0161405P
PR	25-Oct-1939	99US-0161406P
PR	26-Oct-1939	99US-0161535P
PR	26-Oct-1939	99US-0161536P
PR	26-Oct-1939	99US-0161613P
PR	26-Oct-1939	99US-0161970P
PR	28-Oct-1939	99US-0161992P
PR	28-Oct-1939	99US-0161993P
PR	29-Oct-1939	99US-0162142P

Query Match	77.0%	Score 15.4	DB 3	Length 2135
Best Local Similarity	94.1%	Pred. No. 9.1e+02		
Matches 16	Conservative 0	Mismatches 1	Indels 0	Gaps 0

OY		3	TCCATACTCATGTTGAT	19
Dd		39	TCCATACTCATGTTGAT	23

RESULT 32  
AAH54374  
ID AAH54374 standard; DNA; 2946 BP.

AC	AAH54374;
XX	
DT	03-SEP-2001 (first entry)

S. epidermidis genomic polynucleotide sequence SEQ ID NO:3738.

Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination, endocarditis; ds.

OS Staphylococcus epidermidis

PN WO200134809-A2

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US030782

PR 09-NOV-1999; 99US-0164258P

PA (GLAX ) GLAXO GROUP LTD.

PI Kimmerly WJ;

DR WPI; 2001-316495/33.

aa Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*  
 PT useful for vaccinating against infections, e.g. endocarditis.  
 PT

PS Claim 8; Page 1360-1361; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acid (I) encoding polypeptides (II), given in AAG61454 to AAG63120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the S. epidermidis polypeptides (II) via the production of vectors containing them which are used to produce host cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria.





CC obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 28564 BP; 7531 A; 5791 C; 5944 G; 9298 T; 0 U; 0 Other;  
Query Match 77.0%; Score 15.4; DB 9; Length 28564;  
Best Local Similarity 94.1%; Pred. No. 1.3e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
GY 2 GTCCTACTCATGTTGA 18  
DB 19285 GTCCTACTCATGTTGA 19269  
RESULT 37  
ADD48581/C  
ID ADD48581 standard; DNA; 28564 BP.  
AC ADD48581;  
XX  
XX 29-JAN-2004 (first entry)  
DT  
XX Human gene AF297093, SEQ ID NO 14287.  
DE  
XX Human; ds; gene; pain; neuronal tissue; gene therapy;  
XX spinal segmental nerve injury; chronic constriction injury; CCI;  
KM spared nerve injury; SN; Chung.  
XX  
XX Homo sapiens.  
OS  
XX MO2003016475-A2.  
PN  
XX 27-FEB-2003.  
PD  
XX 14-AUG-2002; 2002WO-US025765.  
PF  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-033347P.  
RR  
XX (GEMO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
PI  
XX WPI; 2003-268312/26.  
DR GENBANK; AF297093.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.  
PS  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human DNA (shown in Table 2 of the  
CC specification) which encodes one of the polypeptides of the invention  
CC which is differentially expressed during pain. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 28564 BP; 7531 A; 5791 C; 5944 G; 9298 T; 0 U; 0 Other;  
Query Match 77.0%; Score 15.4; DB 9; Length 28564;  
Best Local Similarity 94.1%; Pred. No. 1.3e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
GY 2 GTCCTACTCATGTTGA 18  
DB 19285 GTCCTACTCATGTTGA 19269  
RESULT 38  
AA161371/C  
ID AA161371 standard; DNA; 335913 BP.  
XX  
XX AA161371;  
AC  
XX 16-OCT-2001 (first entry)  
DT  
XX Soybean 240017 region G3, SEQ ID NO: 2.  
DE  
XX Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;  
XX SCN resistance; rhg4; Rhg4; SCN resistant allele; plant breeding;  
KM 240017 region G3; 318013 region A3; 515002 region G2; ds.  
XX  
XX Glycine max.  
OS  
XX MO200151627-A2.  
PN  
XX 19-JUL-2001.  
PD  
XX 05-JAN-2001; 2001WO-US000552.  
PF  
XX 07-JAN-2000; 2000US-0174880P.  
PR  
XX (MONS ) MONSANTO CO.  
PA  
XX Hauge BM, Wang ML, Parsons JD, Parnell LD;  
PI  
XX WPI; 2001-425872/45.  
DR P-PSDB; AAM42214.  
XX  
XX New purified nucleic acid for producing a soybean plant having soybean  
PT cyst nematode resistance and for use in plant breeding programs.  
XX  
XX Claim 2; Page 204-400; 1353pp; English.  
PS  
XX The invention relates to nucleic acid molecules from regions of the  
CC soybean genome which are associated with soybean cyst nematode (SCN)  
CC resistance. The nucleic acids are used to transform plants, and can  
CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.  
CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes  
CC of soybean plants and for introgressing SCN resistance or partial SCN  
CC resistance into soybean plants. They can also be used in plant breeding  
CC programmes. The invention also relates to proteins encoded by such  
CC nucleic acid molecules, as well as antibodies capable of recognising  
CC these proteins. The present sequence is a nucleic acid molecule provided  
CC in the specification  
XX  
SQ Sequence 335913 BP; 114579 A; 53403 C; 53026 G; 114905 T; 0 U; 0 Other;  
Query Match 77.0%; Score 15.4; DB 5; Length 335913;  
Best Local Similarity 94.1%; Pred. No. 1.7e+03;  
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	3	TCCATACCTCATGTTGAT	19
Db	24565	TCCAAACTCATGTTGAT	24549

  

RESULT 39			
ID	AA161372/C		
XX	AA161372 standard; DNA; 335913 BP.		
XX	AA161372;		
DT	16-OCT-2001 (first entry)		
XX			
DE	Soybean 240017 region G3, SEQ ID NO: 3.		
XX			
KW	Soybean; antihelminthic; gene therapy; soybean cyst nematode; SCN;		
XX	SCN resistance; rhg4; Rhg4; SCN resistant allele; plant breeding;		
KW	240017 region G3; 318013 region A3; 515002 region G2; ds.		
XX			
OS	Glycine max.		
XX			
PN	WO200151627-A2.		
XX			
PD	19-JUL-2001.		
XX			
PF	05-JAN-2001; 2001WO-US000552.		
XX			
PR	07-JAN-2000; 2000US-0174880P.		
XX			
PA	(MONS ) MONSANTO CO.		
XX			
PI	Hauge BM, Wang ML, Parsons JD, Parnell LD;		
XX			
XX	WPI; 2001-425872/45.		
DR	P-PSDB; AAM42215.		
XX			
PT	New purified nucleic acid for producing a soybean plant having soybean		
XX	cyst nematode resistance and for use in plant breeding programs.		
PS	Claim 2; Page 400-595; 1353pp; English.		
XX			
CC	The invention relates to nucleic acid molecules from regions of the		
CC	soybean genome which are associated with soybean cyst nematode (SCN)		
CC	resistance. The nucleic acids are used to transform plants, and can		
CC	produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.		
CC	The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes		
CC	of soybean plants and for introgressing SCN resistance or partial SCN		
CC	resistance into soybean plants. They can also be used in plant breeding		
CC	programmes. The invention also relates to proteins encoded by such		
CC	nucleic acid molecules, as well as antibodies capable of recognising		
CC	these proteins. The present sequence is a nucleic acid molecule provided		
CC	in the specification		
XX			
SEQ	Sequence 335913 BP; 114582 A; 53398 C; 53027 G; 114906 T; 0 U; 0 Other;		

  

QY	3	TCCATACCTCATGTTGAT	19
Db	24565	TCCAAACTCATGTTGAT	24549

  

RESULT 40			
ID	AAAC31848/C		
XX	AAAC31848 standard; cDNA; 308 BP.		
XX	AAAC31848;		
XX			
XX	06-OCT-2000 (first entry)		

```

DE Human secreted protein 5' EST, SEQ ID NO: 35923.
XX
XX
KW Human, 5' EST, expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
PN EPI033401-A2.
PD
PD 06-SEP-2000.
XX
XX
PE 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX
XX (GEST ) GENSET.
XX
XX
XX Dunas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 35923; 71pp + Sequence Listing; English.
XX
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
XX ESTs are derived from mRNAs with intact 5' ends and can therefore be used
XX to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors
XX
XX
XX Sequence 308 BP, 105 A, 54 C, 53 G, 93 T, 0 U, 3 Other;
SQ
OY 1 GGTCCATCACTCATGTTGATG 20
DB 98 GATCCACATCATGTTGATG 79
AC
AC AAAT75448;
XX
XX AAAT75448 standard; DNA, 330 BP.
ID
ID
XX
XX
XX 16-MAR-1999 (first entry)
DT
DE
DE Staphylococcus aureus contig SEQ ID #1137.
XX
XX
XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
XX toxic shock syndrome; ds.
XX
XX
XX Staphylococcus aureus.
XX
XX
XX EPI786519-A2.
XX
XX
XX 30-JUL-1997.

```

XX 07-JAN-1997; 97EP-00100117.  
PF 05-JAN-1996; 96US-0009861P.  
PR (HUMA-) HUMAN GENOME SCI INC.  
PA Kunsh CA, Choi GH, Barash SC, Dillon RJ, Fannon MR, Rosen CA;  
PI WPI; 1997-374922/35.  
XX  
XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus* -  
PT stored on computer readable medium and used in the production of anti-  
PT *S. aureus* vaccines.  
XX  
XX Claim 1; Page 1850; 3271pp; English.  
XX  
XX This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences  
CC of the invention. The DNA sequences are recorded on a computer readable  
CC medium, preferably selected from a floppy or hard disk, random access  
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using  
CC the *S. aureus* DNA sequences allows putative functions to be assigned so  
CC that protein-encoding or regulatory regions of commercial, therapeutic or  
CC industrial importance can be obtained. Specifically, sequences which are  
CC likely to encode antigens have been identified and these polypeptides can  
CC be used in a vaccine composition against *S. aureus* infection. The  
CC polypeptides can also be used in a kit for the immunodetection of  
CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,  
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,  
CC skin and surgical wound infections, scalded skin syndrome, toxic shock  
CC syndrome, etc. Organisms transformed with the DNA sequences can be used  
CC for recombinant production of the polypeptides. The new DNA sequences  
CC (and their fragments) are useful as primers or probes for isolating  
CC homologues of any of the *S. aureus* DNA sequences contained on the computer  
CC readable medium  
SQ Sequence 330 BP; 108 A; 48 C; 54 G; 116 T; 0 U; 4 Other;  
Query Match 76.0%; Score 15.2; DB 2; Length 330;  
Best Local Similarity 85.0%; Pred. No. 8.9e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GGTCCTACTCATGTTGATG 20  
DB 284 GGACCAACTCATGTTGATG 303  
RESULT 42  
AAD50058  
ID AAD50058 standard; DNA; 443 BP.  
XX  
XX AAD50058;  
AC  
XX  
XX 24-MAR-2003 (first entry)  
DT  
XX  
XX Human fatty acid elongation enzyme-like protein DNA #15.  
DE  
XX  
XX Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;  
KW long chain polyunsaturated fatty acid elongation enzyme-like protein;  
KW neurotropic; neuroprotective; antiinflammatory; de.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200264761-A2.  
PN  
XX  
XX 22-AUG-2002.  
PD  
XX  
XX 07-FEB-2002; 2002WO-EP001260.  
PF  
XX  
XX 09-FEB-2001; 2001US-0267415P.  
PR 16-NOV-2001; 2001US-0331449P.  
PR 04-DEC-2001; 2001US-0334948P.  
XX

PA (FARB ) BAYER AG.  
XX  
XX Zhu Z;  
PI  
XX  
XX WPI; 2003-040508/03.  
DR  
XX  
XX Human long chain fatty acid elongation enzyme-like protein is regulated  
PT in the treatment of diseases e.g. cancer, diabetes and CNS disorders.  
PT  
XX  
XX Disclosure; Fig 17, 175pp; English.  
PS  
XX  
XX The invention relates to an isolated polynucleotide encoding a long chain  
CC polyunsaturated fatty acid elongation enzyme-like protein polypeptide.  
CC The invention is useful in the preparation of a medicament for modulating  
CC the activity/function of long chain polyunsaturated fatty acid elongation  
CC enzyme-like protein in a disease specifically cancer, diabetes, a CNS  
CC disorder, metabolic disease, asthma or COPD. The present sequence is  
CC human long chain polyunsaturated fatty acid elongation enzyme-like  
CC protein DNA  
SQ Sequence 443 BP; 170 A; 68 C; 66 G; 139 T; 0 U; 0 Other;  
Query Match 76.0%; Score 15.2; DB 7; Length 443;  
Best Local Similarity 85.0%; Pred. No. 9.3e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GGTCCTACTCATGTTGATG 20  
DB 422 GATCCACATCATGTTGATG 441  
RESULT 43  
AAF11349/c  
ID AAF11349 standard; CDNA; 468 BP.  
XX  
XX AAF11349;  
AC  
XX  
XX 13-MAR-2001 (first entry)  
DT  
XX  
XX *Aspergillus niger* EST SEQ ID NO:3872.  
DE  
XX  
XX Multiple gene expression; filamentous fungal cell; EST;  
KW expressed sequence tag; *Fusarium venenatum*; *Aspergillus niger*;  
KW *Aspergillus oryzae*; *Trichoderma reesei*; identification; recombination;  
KW culture condition; environmental stress; spore morphogenesis;  
KW metabolic pathway engineering; catalytic pathway engineering; ss.  
XX  
XX *Aspergillus niger*.  
OS  
XX  
XX WO200056762-A2.  
PN  
XX  
XX 28-SEP-2000.  
PD  
XX  
XX 22-MAR-2000; 2000WO-US007781.  
PF  
XX  
XX 22-MAR-1999; 99US-00273623.  
PR  
XX  
XX (NOVO ) NOVO NORDISK BIOTECH INC.  
PA (NOVO ) NOVO NORDISK AS.  
PA  
XX  
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
PI WPI; 2000-594572/56.  
DR  
XX  
XX Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags.  
XX  
XX Claim 87; Page 1730; 3161pp; English.  
PS  
XX  
XX The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal





CC polypeptides can also be used to aid wound healing and epithelial cell  
CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
CC before transplantation, for supporting cell culture of primary tissues,  
CC to regenerate tissues and in chemotaxis. The polypeptides can also be  
CC used as a food additive or preservative to increase or decrease storage  
CC capabilities. ABK69078-ABK69143 represent human secreted protein coding  
CC sequences, PCR primers and related sequences used in cloning and  
CC expression of the secreted proteins described in examples of the  
CC invention  
XX  
SQ Sequence 479 BP, 106 A, 169 C, 107 G, 97 T, 0 U, 0 Other;  
  
Query Match 76.0%; Score 15.2; DB 6; Length 479;  
Best Local Similarity 85.0%; Pred. No. 9.4e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 GGTCGACTCATGTTGATG 20  
DB 373 GGTCGACTCATGTTGATG 354  
  
RESULT 46  
AAD50057/c  
ID AAD50057 standard; DNA; 499 BP.  
XX  
AC AAD50057;  
XX  
DT 24-MAR-2003 (first entry)  
XX  
DE Human fatty acid elongation enzyme-like protein DNA #14.  
XX  
KW Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;  
KW long chain polyunsaturated fatty acid elongation enzyme-like protein;  
KW neurotropic; neuroprotective; antiinflammatory; de.  
XX  
OS Homo sapiens.  
XX  
PN WO200264761-A2.  
XX  
PD 22-AUG-2002.  
XX  
PF 07-FEB-2002; 2002WO-EP001260.  
XX  
PR 09-FEB-2001; 2001US-0267415P.  
PR 16-NOV-2001; 2001US-0331449P.  
PR 04-DEC-2001; 2001US-0334948P.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Zhu Z;  
XX  
DR WPI; 2003-040508/03.  
XX  
PT Human long chain fatty acid elongation enzyme-like protein is regulated  
PT in the treatment of diseases e.g. cancer, diabetes and CNS disorders.  
XX  
PS Disclosure; Fig 16; 175pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a long chain  
CC polyunsaturated fatty acid elongation enzyme-like protein polypeptide.  
CC The invention is useful in the preparation of a medicament for modulating  
CC the activity/function of long chain polyunsaturated fatty acid elongation  
CC enzyme-like protein in a disease specifically cancer, diabetes, a CNS  
CC disorder, metabolic disease, asthma or COPD. The present sequence is  
CC human long chain polyunsaturated fatty acid elongation enzyme-like  
CC protein DNA  
XX  
SQ Sequence 499 BP, 174 A, 97 C, 82 G, 145 T, 0 U, 1 Other;  
  
Query Match 76.0%; Score 15.2; DB 7; Length 499;  
Best Local Similarity 85.0%; Pred. No. 9.4e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCGACTCATGTTGATG 20  
DB 304 GATCCACATCATGTTGATG 285  
  
RESULT 47  
ABK69131/c  
ID ABK69131 standard; cDNA; 547 BP.  
XX  
AC ABK69131;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE DNA encoding human secreted protein, SEQ ID No 55.  
XX  
KW Human; secreted protein; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm; breast; liver; ischaemia;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;  
KW epithelial cell proliferation; organ transplantation; food additive;  
KW food storage; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200224721-A1.  
XX  
PD 28-MAR-2002.  
XX  
PF 09-JAN-2001; 2001WO-US000544.  
XX  
PR 20-SEP-2000; 2000US-0234211P.  
XX  
PA (HDNA-) HUMAN GENOME SCT INC.  
XX  
PI Komatsu, GA, Baker KP, Rosen CA, Birse CE, Soppet DR;  
PI Olsen HS, Ni J, Fiscella M, Moore PA, Wei P, Ebner R, Duan DR;  
PI Shi Y, Choi GH;  
XX  
DR WPI; 2002-330012/36.  
DR P-PSDB; AAU96210.  
XX  
PT Isolated nucleic acid molecule encoding a human secreted protein is used  
PT in preventing, treating or ameliorating a medical condition.  
XX  
PS Claim 1; Page 484; 562pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) encoding a  
CC human secreted protein (II). (I) and (II) are used to prevent, treat or  
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,  
CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in  
CC diagnosing a pathological condition or susceptibility to a pathological  
CC condition. The antibodies to (II) can also be used in alleviating  
CC symptoms associated with the disorders and in diagnostic immunoassays  
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).  
CC Disorders which are diagnosed or treated include autoimmune diseases e.g.  
CC rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the  
CC breast or liver, cardiovascular disorders e.g. cardiac arrest,  
CC cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous  
CC system disorders e.g. Alzheimer's disease, infections caused by bacteria,  
CC viruses and fungi and ocular disorders e.g. corneal infection. The  
CC polypeptides can also be used to aid wound healing and epithelial cell  
CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
CC before transplantation, for supporting cell culture of primary tissues,  
CC to regenerate tissues and in chemotaxis. The polypeptides can also be  
CC used as a food additive or preservative to increase or decrease storage  
CC capabilities. ABK69078-ABK69143 represent human secreted protein coding  
CC sequences, PCR primers and related sequences used in cloning and  
CC expression of the secreted proteins described in examples of the  
CC invention  
XX  
SQ Sequence 547 BP, 119 A, 185 C, 121 G, 117 T, 0 U, 5 Other;

```
Query Match          76.0%; Score 15.2; DB 6; Length 547;
Best Local Similarity 85.0%; Pred. No. 9.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GGTCCATCTCATGTTGATG 20
        |||||
DB      366 GGTCCATCTCATGTTGATG 347

RESULT 48
AAD50054
ID      AAD50054 standard; DNA; 548 BP.
XX
AC      AAD50054;
XX
DT      24-MAR-2003 (first entry)
XX
DE      Human fatty acid elongation enzyme-like protein DNA #11.
XX
KW      Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;
KM      long chain polyunsaturated fatty acid elongation enzyme-like protein;
KW      neurotropic; neuroprotective; antiinflammatory; ds.
XX
OS      Homo sapiens.
XX
PN      WO200264761-A2.
XX
PD      22-AUG-2002.
XX
PF      07-FEB-2002; 2002WO-EP001260.
XX
PR      09-FEB-2001; 2001US-0267415P.
XX
PR      16-NOV-2001; 2001US-0331449P.
XX
PR      04-DEC-2001; 2001US-0334948P.
XX
PA      (FARB ) BAYER AG.
XX
PI      Zhu Z;
XX
DR      WPI; 2003-040508/03.
XX
PT      Human long chain fatty acid elongation enzyme-like protein is regulated
PT      in the treatment of diseases e.g. cancer, diabetes and CNS disorders.
XX
PS      Disclosure; Fig 13; 175pp; English.
XX
CC      The invention relates to an isolated polynucleotide encoding a long chain
CC      polyunsaturated fatty acid elongation enzyme-like protein polypeptide.
CC      The invention is useful in the preparation of a medicament for modulating
CC      the activity/function of long chain polyunsaturated fatty acid elongation
CC      enzyme-like protein in a disease specifically cancer, diabetes, a CNS
CC      disorder, metabolic disease, asthma or COPD. The present sequence is
CC      human long chain polyunsaturated fatty acid elongation enzyme-like
CC      protein DNA
XX
SQ      Sequence 548 BP; 192 A; 84 C; 86 G; 186 T; 0 U; 0 Other;

Query Match          76.0%; Score 15.2; DB 7; Length 548;
Best Local Similarity 85.0%; Pred. No. 9.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GGTCCATCTCATGTTGATG 20
        |||||
DB      422 GATCCACATCATGTTGATG 441

RESULT 49
AAD50053
ID      AAD50053 standard; DNA; 578 BP.
XX
AC      AAD50053;
XX
DT      24-MAR-2003 (first entry)
```

```
XX      Human fatty acid elongation enzyme-like protein DNA #10.
DE      Human; cancer; diabetes; CNS disorder; metabolic disease; asthma; COPD;
XX      long chain polyunsaturated fatty acid elongation enzyme-like protein;
KM      neurotropic; neuroprotective; antiinflammatory; ds.
XX
OS      Homo sapiens.
XX
PN      WO200264761-A2.
XX
PD      22-AUG-2002.
XX
PF      07-FEB-2002; 2002WO-EP001260.
XX
PR      09-FEB-2001; 2001US-0267415P.
XX
PR      16-NOV-2001; 2001US-0331449P.
XX
PR      04-DEC-2001; 2001US-0334948P.
XX
PA      (FARB ) BAYER AG.
XX
PI      Zhu Z;
XX
DR      WPI; 2003-040508/03.
XX
PT      Human long chain fatty acid elongation enzyme-like protein is regulated
PT      in the treatment of diseases e.g. cancer, diabetes and CNS disorders.
XX
PS      Disclosure; Fig 12; 175pp; English.
XX
CC      The invention relates to an isolated polynucleotide encoding a long chain
CC      polyunsaturated fatty acid elongation enzyme-like protein polypeptide.
CC      The invention is useful in the preparation of a medicament for modulating
CC      the activity/function of long chain polyunsaturated fatty acid elongation
CC      enzyme-like protein in a disease specifically cancer, diabetes, a CNS
CC      disorder, metabolic disease, asthma or COPD. The present sequence is
CC      human long chain polyunsaturated fatty acid elongation enzyme-like
CC      protein DNA
XX
SQ      Sequence 578 BP; 199 A; 87 C; 87 G; 205 T; 0 U; 0 Other;

Query Match          76.0%; Score 15.2; DB 7; Length 578;
Best Local Similarity 85.0%; Pred. No. 9.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GGTCCATCTCATGTTGATG 20
        |||||
DB      439 GATCCACATCATGTTGATG 458

RESULT 50
AAF14157
ID      AAF14157 standard; cDNA; 685 BP.
XX
AC      AAF14157;
XX
DT      13-MAR-2001 (first entry)
XX
DE      Aspergillus oryzae EST SEQ ID NO:6680.
XX
KW      Multiple gene expression; filamentous fungal cell; EST;
KW      expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW      Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW      culture condition; environmental stress; spore morphogenesis;
KW      metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS      Aspergillus oryzae.
XX
PN      WO200056762-A2.
XX
PD      28-SEP-2000.
XX
PF      22-MAR-2000; 2000WO-US007781.
```

XX 22-MAR-1999; 99US-00273623.

XX (NOVO ) NOVO NORDISK BIOTECH INC.  
PA (NOVO ) NOVO NORDISK AS.

XX Berka RM, Rey MM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;  
PI WPI; 2000-594572/56.

XX Monitoring differential expression of genes in filamentous fungal cells  
PT uses fluorescence-labeled nucleic acids isolated from the cells and a  
PT substrate of expressed sequence tags.

PS Claim 88; Page 2728-2729; 3161pp; English.

XX The present invention describes a method for monitoring differential  
CC expression of genes in a first filamentous fungal (FF) cell relative to  
CC expression of the same genes in one or more second filamentous fungal  
CC cells. The method uses fluorescence-labeled nucleic acids isolated from  
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs  
CC are used in the methods for monitoring differential expression of genes  
CC in a first filamentous fungal (FF) cell relative to expression of the  
CC same genes in one or more second filamentous fungal cells. Monitoring the  
CC global expression of genes from FF cells allows the production potential  
CC of the microorganisms to be improved. New genes may be discovered,  
CC possible functions of unknown open reading frames can be identified and  
CC gene copy number variation and stability can be monitored. The expression  
CC of genes can be used to study how FF cells adapt to changes in culture  
CC conditions, environmental stress, spore morphogenesis, recombination,  
CC metabolic or catabolic pathway engineering. Using ESTs provides several  
CC advantages over genomic or random cDNA clones including elimination of  
CC redundancy as one spot on an array equals one gene or open reading frame,  
CC and organization of the microarrays based on function of the gene  
CC products to facilitate analysis of the results. AAF07478 to AAF11247  
CC represents ESTs from Fusarium venenatum; AAF11248 to AAF11853 represents  
CC ESTs from Aspergillus niger; AAF11854 to AAF14878 represents ESTs from  
CC Aspergillus oryzae; and AAF14879 to AAF15337 represents ESTs from  
CC Trichoderma reesei, which are all specifically claimed in the present  
CC invention

SQ Sequence 685 BP; 169 A; 178 C; 174 G; 164 T; 0 U; 0 Other;

Query Match 76.0%; Score 15.2; DB 3; Length 685;  
Best Local Similarity 85.0%; Pred. No. 9.8e+02;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGATG 20  
|||  
Db 353 GGTCATCTCTGTGATGCTG 372

Search completed: August 17, 2004, 14:36:28  
Job time : 402 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2004, 14:16:42 ; Search time 2541 Seconds  
(without alignments)  
235.043 Million cell updates/sec

Title: US-09-825-489-3

Sequence: 1 ggtccatcatcatgtcatg 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 800 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estinu:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inu:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
20	100.0	219	12	BG201597	BG201597 RST20938
20	100.0	221	12	BG212172	BG212172 RST1643
20	100.0	352	12	BG185555	BG185555 RST4504 A
20	100.0	359	14	CB852193	CB852193 UI-CF-FNO

5	100.0	370	12	BM796479	BM796479 K-EST0079
6	100.0	408	9	AV687532	AV687532 AV687532
7	100.0	408	9	AV690967	AV690967 AV690967
8	100.0	437	9	AI961077	AI961077 wg44g01.x
9	100.0	461	13	BM679004	BM679004 UI-CF-DU1
10	100.0	486	9	AA628925	AA628925 od81a10.8
11	100.0	487	12	BI793276	BI793276 1e50b09.y
12	100.0	506	9	AA167098	AA167098 zpo5g10.x
13	100.0	510	13	BQ186624	BQ186624 UI-E-EJ1-
14	100.0	513	9	AA165911	AA165911 zp05g10.8
15	100.0	515	13	BQ184648	BQ184648 UI-E-EJ1-
16	100.0	534	10	AM890854	AM890854 RCS-NT005
17	100.0	587	14	CB048344	CB048344 NISC_9104
18	100.0	591	14	CB048343	CB048343 NISC_9104
19	100.0	592	10	AM979099	AM979099 EST391209
20	100.0	595	12	BM739320	BM739320 K-EST0008
21	100.0	598	12	BI792983	BI792983 1e50b09.x
22	100.0	627	10	BF446397	BF446397 7p36g05.x
23	100.0	638	12	BM853017	BM853017 K-EST0134
24	100.0	639	12	BI911749	BI911749 603065343
25	100.0	672	14	CB305774	CB305774 UI-CF-PN1
26	100.0	678	12	BM929090	BM929090 UI-H-DL1-
27	100.0	705	14	CA450156	CA450156 UI-CF-FNO
28	100.0	748	13	BQ772342	BQ772342 UI-H-EZ1-
29	100.0	790	12	BQ186679	BQ186679 RST5654 A
30	100.0	809	29	AY414079	AY414079 Pan trogl
31	100.0	813	29	AY414078	AY414078 Homo sapi
32	100.0	873	9	AL545429	AL545429 AL545429
33	100.0	1059	12	BM468458	BM468458 AGNCNCOURT
34	100.0	1114	9	AL573331	AL573331 AL573331
35	100.0	1129	9	AL548961	AL548961 AL548961
36	100.0	1174	13	BQ879392	BQ879392 AGNCNCOURT
37	98.0	1201	9	AL531628	AL531628 AL531628
38	19	640	10	BE763751	BE763751 RCS-NT005
39	92.0	568	13	BQ321319	BQ321319 QVA-CF049
40	18.4	625	12	BQ057361	BQ057361 BQ057361
41	92.0	711	9	AT634902	AT634902 t274c02.x
42	18.4	854	12	BQ187563	BQ187563 RST6691 A
43	92.0	857	9	AL571156	AL571156 AL571156
44	18.4	1024	28	CC448351	CC448351 ZMMBNC032
45	18.4	420	10	AM646323	AM646323 cm63e10.w
46	17.4	492	12	BQ083380	BQ083380 BQ083380
47	87.0	587	9	AL637062	AL637062 AL637062
48	17.4	828	14	CD434346	CD434346 EL01N0323
49	17.4	85.0	12	BI261004	BI261004 602972177
50	17	844	29	CG932703	CG932703 MBHEW59TF
51	85.0	198	29	CG486388	CG486388 OSTR1372
52	16.8	294	12	BI297979	BI297979 UI-R-CV2-
53	16.8	348	12	BI297518	BI297518 UI-R-CV2-
54	16.8	360	12	BM293462	BM293462 C0304G12-
55	16.8	365	9	AA927997	AA927997 on79h12.8
56	16.8	390	12	BQ031452	BQ031452 UI-1-CFO-
57	16.8	436	10	BF417305	BF417305 UI-R-CNO-
58	16.8	500	12	BI300394	BI300394 UI-R-CV2-
59	16.8	518	28	CG638173	CG638173 OSTR67192
60	16.8	531	29	CG112083	CG112083 PUIFX20TD
61	16.8	563	14	CA351074	CA351074 622055 NC
62	16.8	597	13	BM243041	BM243041 BM243041
63	16.8	598	13	BQ207028	BQ207028 UI-R-DZ1-
64	16.8	616	14	CA381609	CA381609 661145 NC
65	16.8	632	14	CD292702	CD292702 STRPUS36.
66	16.8	635	14	CA379904	CA379904 659105 NC
67	16.8	636	9	AV864708	AV864708 AV864708
68	16.8	657	29	AG115048	AG115048 Pan trogl
69	16.8	667	10	BF155636	BF155636 7m92907.x
70	16.8	669	13	BM261667	BM261667 BM261667
71	16.8	679	14	CA231933	CA231933 SCQSLF303
72	16.8	681	13	BM270720	BM270720 BM270720
73	16.8	691	13	EX075094	EX075094 BX075094
74	16.8	700	14	CA356753	CA356753 629000 NC
75	16.8	722	13	CA060694	CA060694 sea1rgb53
76	16.8				
77	16.8				

78	16.8	84.0	730	12	BI296724	UI-R-CV2-	C 151	15.8	79.0	404	9	A1446339	A1446339 tJ20905.x
79	16.8	84.0	757	14	CB511930	ssalrpb54	C 152	15.8	79.0	411	9	A1624772	A1624772 t644c12.x
80	16.8	84.0	789	28	BZ982973	PUGCR07TB	C 153	15.8	79.0	413	9	AL544616	AL544616 AL544616
81	16.8	84.0	811	29	CG555961	OGMCH90TV	C 154	15.8	79.0	415	10	H05220	H05220 Y185d09..81
82	16.8	84.0	819	29	AXY414080	Mus muscu	C 155	15.8	79.0	419	10	AM273418	AM273418 xc39c09.x
83	16.8	84.0	830	28	BH710049	BOMOU27TR	C 156	15.8	79.0	443	13	BO973507	BO973507 OH113A16
84	16.8	84.0	853	28	BH086317	RPCT-24-2	C 157	15.8	79.0	444	9	A1490900	A1490900 EST7213609
85	16.8	84.0	870	13	BUB40815	RPCT-24-2	C 158	15.8	79.0	444	9	A1564466	A1564466 t657f03.x
86	16.8	84.0	984	29	CNS006KD	AL077281	C 159	15.8	79.0	450	10	BF894540	BF894540 RCI-WT013
87	16.8	84.0	1682	10	BF572295	Drosoph11	C 160	15.8	79.0	451	28	AO917107	AO917107 T233136b
88	16.4	82.0	295	10	BB286508	BF572295	C 161	15.8	79.0	452	14	N25249	N25249 YX74d06..81
89	16.4	82.0	363	14	CAS54445	BB286508	C 162	15.8	79.0	453	9	AA805564	AA805564 oc24d07.s
90	16.4	82.0	400	28	AO011724	CAS54445	C 163	15.8	79.0	456	10	AM195941	AM195941 X185a12.x
91	16.4	82.0	400	28	AO085335	AO011724	C 164	15.8	79.0	457	12	CD283195	CD283195 G38941.42
92	16.4	82.0	405	28	BH145400	AO085335	C 165	15.8	79.0	461	28	BH883167	BH883167 hw42f04..9
93	16.4	82.0	513	28	BH660918	BH145400	C 166	15.8	79.0	469	10	BF435867	BF435867 nab24n11..
94	16.4	82.0	522	29	BK652083	BH660918	C 167	15.8	79.0	470	14	H02898	H02898 yJ39e03..81
95	16.4	82.0	602	29	CE524885	BK652083	C 168	15.8	79.0	471	9	AA743710	AA743710 ny92f04..8
96	16.4	82.0	620	14	CA548088	CE524885	C 169	15.8	79.0	475	12	B1582117	B1582117 Rk20037.5
97	16.4	82.0	626	28	AO286992	CA548088	C 170	15.8	79.0	477	10	AM570406	AM570406 gJ22h03..Y
98	16.4	82.0	668	28	AZ333406	AO286992	C 171	15.8	79.0	477	28	AO814813	AO814813 HS..5258..B
99	16.4	82.0	672	10	BB053667	AZ333406	C 172	15.8	79.0	479	9	AA725575	AA725575 a124d11..8
100	16.4	82.0	702	28	AZ434015	BB053667	C 173	15.8	79.0	481	10	BE074235	BE074235 QV3-BT056
101	16.4	82.0	775	14	CA429356	AZ434015	C 174	15.8	79.0	483	28	AO689450	AO689450 tpxxb0079N
102	16.4	82.0	775	28	BH483076	CA429356	C 175	15.8	79.0	485	12	BM144489	BM144489 TCAPAPD11
103	16.4	82.0	779	29	BX138283	BH483076	C 176	15.8	79.0	486	12	BM769521	BM769521 K-EST052
104	16.4	82.0	805	28	BH363463	BX138283	C 177	15.8	79.0	487	12	B1786529	B1786529 ba150b09..
105	16.4	82.0	824	28	AZ174590	BH363463	C 178	15.8	79.0	507	9	A1889130	A1889130 tw36h12..x
106	16.4	82.0	831	29	CC511858	AZ174590	C 179	15.8	79.0	508	14	CA735557	CA735557 wP18..PK0
107	16.4	82.0	914	13	BU074772	CC511858	C 180	15.8	79.0	512	28	AZ698834	AZ698834 RPCI-23-2
108	16.4	82.0	958	28	AO743158	BU074772	C 181	15.8	79.0	512	14	CB085136	CB085136 h427d12..b
109	16.4	82.0	1002	10	BE616227	AO743158	C 182	15.8	79.0	514	28	B2916626	B2916626 CH240..60U
110	16.4	82.0	2320	11	AK032870	BE616227	C 183	15.8	79.0	516	10	BE835251	BE835251 RC5-FM002
111	16.4	82.0	3253	11	AK030322	AK032870	C 184	15.8	79.0	518	10	BE835251	BE835251 RC5-FM002
112	16.4	82.0	480	14	CP931227	AK030322	C 185	15.8	79.0	523	10	BZ221678	BZ221678 XB83c03..x
113	16.4	82.0	575	28	AZ019301	CP931227	C 186	15.8	79.0	526	28	BZ221678	BZ221678 XB83c03..x
114	16.4	82.0	576	12	BU053329	AZ019301	C 187	15.8	79.0	527	13	BU713949	BU713949 SJAABXA07
115	16.4	82.0	636	10	BF936583	BU053329	C 188	15.8	79.0	533	13	BU950227	BU950227 1o78c09..x
116	16.4	82.0	863	29	CG857171	BF936583	C 189	15.8	79.0	537	12	B1887105	B1887105 ZF637-1-0
117	15.8	79.0	204	10	BE243298	CG857171	C 190	15.8	79.0	541	9	AV914909	AV914909 AV914909
118	15.8	79.0	207	14	H46770	BE243298	C 191	15.8	79.0	545	13	BU927775	BU927775 AGENCOURT
119	15.8	79.0	264	12	BM151116	H46770	C 192	15.8	79.0	548	12	B1498599	B1498599 BA116C04..
120	15.8	79.0	270	14	CB496164	BM151116	C 193	15.8	79.0	552	9	A1957898	A1957898 fd07g05..x
121	15.8	79.0	270	28	BZ215304	CB496164	C 194	15.8	79.0	554	12	BG524360	BG524360 41-72..Ste
122	15.8	79.0	273	10	AM367203	BZ215304	C 195	15.8	79.0	557	9	AL045488	AL045488 DKF2P4341
123	15.8	79.0	275	9	AA665903	AM367203	C 196	15.8	79.0	560	28	AO690181	AO690181 tpxxb0081D
124	15.8	79.0	275	10	AM196133	AA665903	C 197	15.8	79.0	564	10	AM469256	AM469256 hc80g05..x
125	15.8	79.0	278	13	BO630957	AM196133	C 198	15.8	79.0	564	14	CP032818	CP032818 QCB8e02..Y
126	15.8	79.0	297	14	CF087871	BO630957	C 199	15.8	79.0	574	10	BF382660	BF382660 60t816872
127	15.8	79.0	297	14	R24247	CF087871	C 200	15.8	79.0	574	29	CE756403	CE756403 t1gr-g8s-
128	15.8	79.0	299	28	CC015488	R24247	C 201	15.8	79.0	579	29	CE439041	CE439041 t1gr-g8s-
129	15.8	79.0	313	12	BMS10485	CC015488	C 202	15.8	79.0	584	10	BF648403	BF648403 NFO47B08E
130	15.8	79.0	314	12	BO582223	BMS10485	C 203	15.8	79.0	587	13	EX487241	EX487241 DKF2P5860
131	15.8	79.0	317	13	BMS11316	BO582223	C 204	15.8	79.0	589	28	AO986586	AO986586 RPCI-23-3
132	15.8	79.0	327	13	BU918258	BMS11316	C 205	15.8	79.0	590	28	AZ421903	AZ421903 IM0200C14
133	15.8	79.0	329	14	R91378	BU918258	C 206	15.8	79.0	592	14	CB163502	CB163502 K-BST0224
134	15.8	79.0	343	12	B1608116	R91378	C 207	15.8	79.0	593	28	AZ406088	AZ406088 IM0175M09
135	15.8	79.0	347	14	N20854	B1608116	C 208	15.8	79.0	599	10	AM673731	AM673731 ba57c01..x
136	15.8	79.0	358	14	D62710	N20854	C 209	15.8	79.0	599	28	AO438485	AO438485 HS..5115..B
137	15.8	79.0	362	10	AM651621	D62710	C 210	15.8	79.0	600	28	BZ371495	BZ371495 i657b12..g
138	15.8	79.0	362	29	CG943525	AM651621	C 211	15.8	79.0	600	28	BZ375237	BZ375237 i657b12..g
139	15.8	79.0	365	28	AO015910	CG943525	C 212	15.8	79.0	605	28	AZ257832	AZ257832 RPCI-23-1
140	15.8	79.0	368	9	AL080028	AO015910	C 213	15.8	79.0	611	14	CB483863	CB483863 cc1uwbh0
141	15.8	79.0	368	10	AM512370	AL080028	C 214	15.8	79.0	615	28	B74016	B74016 t27h20TR..TA
142	15.8	79.0	383	12	BM782600	AM512370	C 215	15.8	79.0	620	14	CP307533	CP307533 HD11--06-
143	15.8	79.0	384	29	CC805486	BM782600	C 216	15.8	79.0	622	14	CP807709	CP807709 p8HB027XI
144	15.8	79.0	388	9	A1539027	CC805486	C 217	15.8	79.0	631	13	BQ410891	BQ410891 GA..Ed003
145	15.8	79.0	388	12	BI055583	A1539027	C 218	15.8	79.0	635	14	CD014962	CD014962 hac27A10..
146	15.8	79.0	389	14	H41924	BI055583	C 219	15.8	79.0	638	28	AO602236	AO602236 HS..2111..B
147	15.8	79.0	391	14	N20864	H41924	C 220	15.8	79.0	645	14	CB688581	CB688581 CEST-27-D
148	15.8	79.0	395	14	CF186400	N20864	C 221	15.8	79.0	653	13	BY722968	BY722968 BY722968
149	15.8	79.0	398	12	BI074444	CF186400	C 222	15.8	79.0	658	28	AZ282626	AZ282626 RPCI-23-1
150	15.8	79.0	398	12	BI074444	BI074444	C 223	15.8	79.0	661	28	BZ006698	BZ006698 oen96a02..

C 224	15.8	79.0	666	29	CG760052	ZMMBBD031	C 297	15.4	77.0	352	12	BI771144	BI771144	EBrc003	SQ	
C 225	15.8	79.0	668	14	CD079743	MA3_-99990	298	15.4	77.0	365	28	A2255310	A2255310	RPCI_-23_-9		
C 226	15.8	79.0	673	14	CB445568	666932	MA	C 299	15.4	77.0	372	9	AA396945	AA396945	mr1f10.x	
C 227	15.8	79.0	673	14	CB449974	704590	MA	C 300	15.4	77.0	376	9	AI851749	AI851749	UI-M-BH0-	
C 228	15.8	79.0	675	29	CG855018	ZMMBBD022		C 301	15.4	77.0	380	14	CF506831	CF506831	USDA-EP_1	
C 229	15.8	79.0	676	12	BM468848	AGENCOCRT		C 302	15.4	77.0	381	9	AI216097	AI216097	gm40b03.x	
C 230	15.8	79.0	679	29	CE661268	l1gr-g8s-		C 303	15.4	77.0	386	12	BG232692	BG232692	Pa78g04.y	
C 231	15.8	79.0	695	28	BH506396	BH506396	BOGV61TF	C 304	15.4	77.0	392	13	BY605976	BY605976	BY605976	
C 232	15.8	79.0	699	9	AV938999	AV938999		C 305	15.4	77.0	393	9	BI766819	BI766819	ah37a12.s	
C 233	15.8	79.0	699	14	CB688862	CEST_-26-B		C 306	15.4	77.0	393	9	AA209714	AA209714	mw72b02.x	
C 234	15.8	79.0	704	12	BI931092	EST550981		C 307	15.4	77.0	400	14	CF506703	CF506703	USDA-FP_1	
C 235	15.8	79.0	707	13	BM038685	BM038685		C 308	15.4	77.0	407	9	AI655581	AI655581	lc27h12.x	
C 236	15.8	79.0	711	14	CA423177	UI-H-FE1-		C 309	15.4	77.0	407	10	BF061808	BF061808	7k67C02.x	
C 237	15.8	79.0	714	14	CA413785	UI-H-EZ0-		C 310	15.4	77.0	407	13	BY228504	BY228504	BY228504	
C 238	15.8	79.0	714	28	BH578920	BOCON57FP		C 311	15.4	77.0	421	10	R49177	R49177		
C 239	15.8	79.0	716	28	BH941998	odf46h06.		C 312	15.4	77.0	428	29	EX189233	EX189233		
C 240	15.8	79.0	722	12	BIJ51331	BIJ51331		C 313	15.4	77.0	429	9	AA387236	AA387236	Danio_rer	
C 241	15.8	79.0	722	12	BC468373	60250983		C 314	15.4	77.0	435	10	BE944649	BE944649	vc21g11.x	
C 242	15.8	79.0	728	10	BE381698	601272121		C 315	15.4	77.0	441	9	AA013929	AA013929	UI-M-BH3-	
C 243	15.8	79.0	736	29	CE417738	l1gr-g8s-		C 316	15.4	77.0	449	13	BI113931	BI113931	mh24g08.x	
C 244	15.8	79.0	738	28	BZ523914	BORAF75TF		C 317	15.4	77.0	450	14	CD372378	CD372378	UI-R-GD0-	
C 245	15.8	79.0	739	12	BG615504	602642809		C 318	15.4	77.0	452	9	AA881328	AA881328	vx12h08.x	
C 246	15.8	79.0	742	28	BH650112	BOMBL43TF		C 319	15.4	77.0	453	28	AQ0117439	AQ0117439	AA0117439	
C 247	15.8	79.0	747	29	CE567553	l1gr-g8s-		C 320	15.4	77.0	455	13	BY465273	BY465273	BY465273	
C 248	15.8	79.0	753	14	CD743096	UI-H-Ft1-		C 321	15.4	77.0	456	9	AA013931	AA013931	mh24h08.x	
C 249	15.8	79.0	755	14	CF738663	UI-M-HD0-		C 322	15.4	77.0	474	10	BF660500	BF660500	ma92f12.	
C 250	15.8	79.0	767	14	CB229722	AGENCOCRT		C 323	15.4	77.0	476	14	CF156184	CF156184	CF156184	
C 251	15.8	79.0	767	28	BH650336	BH650336	BOHZY44TR	C 324	15.4	77.0	477	14	CF158701	CF158701	B0663A01-	
C 252	15.8	79.0	769	28	BH490051	BOHGU57TR		C 325	15.4	77.0	477	14	CF158761	CF158761	CF158761	
C 253	15.8	79.0	783	28	BH435182	BH435182	BOHAZ61TR	C 326	15.4	77.0	478	10	BF657849	BF657849	ma92f12.	
C 254	15.8	79.0	787	14	CA306138	CA306138		C 327	15.4	77.0	478	14	CA555095	CA555095	CO893CG09-	
C 255	15.8	79.0	796	9	AL580875	AL580875		C 328	15.4	77.0	481	13	BY243831	BY243831	BY243831	
C 256	15.8	79.0	800	12	BI932916	EST552805		C 329	15.4	77.0	486	14	CF155910	CF155910	B0630P09-	
C 257	15.8	79.0	807	29	CG369334	OG0G026TF		C 330	15.4	77.0	486	28	BH661806	BH661806	BOMF128TR	
C 258	15.8	79.0	809	28	BH696645	BOMF181TF		C 331	15.4	77.0	490	28	CG325458	CG325458	TEA105_Ba	
C 259	15.8	79.0	834	9	AL529131	AL529131		C 332	15.4	77.0	492	28	BH385412	BH385412	AG-ND_-131	
C 260	15.8	79.0	851	10	BF692242	602249113		C 333	15.4	77.0	496	10	AW682703	AW682703	EST00976	
C 261	15.8	79.0	856	28	BZ447974	BONF31TR		C 334	15.4	77.0	510	12	BG524561	BG524561	43_-B3_Ste	
C 262	15.8	79.0	863	28	B20944	TBE3-T7.2.T		C 335	15.4	77.0	515	14	CD291038	CD291038	St-FPu538.	
C 263	15.8	79.0	865	14	CK291457	EST754171		C 336	15.4	77.0	530	12	BM403834	BM403834	EST758161	
C 264	15.8	79.0	872	29	CG369341	OG0G026TF		C 337	15.4	77.0	530	14	CD290266	CD290266	St-FPu538.	
C 265	15.8	79.0	881	28	CG388186	PUNHOX4TB		C 338	15.4	77.0	531	29	CG427151	CG427151	0180726-0	
C 266	15.8	79.0	901	12	BG622103	BG622103	602646717	C 339	15.4	77.0	534	10	BE404894	BE404894	WHE1206.C	
C 267	15.8	79.0	913	9	AL520808	AL520808		C 340	15.4	77.0	536	29	CC799458	CC799458	0180559-0	
C 268	15.8	79.0	921	12	BM358157	GA_Ea000		C 341	15.4	77.0	542	10	BF466844	BF466844	UI-M-GCOP	
C 269	15.8	79.0	935	10	BF181934	601807280		C 342	15.4	77.0	555	12	BG520816	BG520816	ps52c05.y	
C 270	15.8	79.0	938	13	BX328627	BX328627		C 343	15.4	77.0	556	10	BF427832	BF427832	dE91g08.y	
C 271	15.8	79.0	966	13	BX328628	BX328628		C 344	15.4	77.0	565	28	B2173300	B2173300	CH230_-321	
C 272	15.8	79.0	979	28	B08657	B08657	T1B10-T7_TA	C 345	15.4	77.0	569	12	BM813690	BM813690	EST591783	
C 273	15.8	79.0	1011	13	BQ706318	BQ706318	AGENCOCRT	C 346	15.4	77.0	574	28	AQ524050	AQ524050	HS_-5232_B	
C 274	15.8	79.0	1060	9	AL575835	AL575835		C 347	15.4	77.0	585	10	BE775870	BE775870	MY_-08-A-0	
C 275	15.8	79.0	1096	12	BG329874	602429558		C 348	15.4	77.0	593	14	BF730492	BF730492	ma6b4A06.	
C 276	15.8	79.0	1101	28	CC206155	CH261-138		C 349	15.4	77.0	594	14	CD311937	CD311937	St-FPu691.	
C 277	15.8	79.0	1125	10	BF790487	602250157		C 350	15.4	77.0	597	12	BI414455	BI414455	St-FPu691.	
C 278	15.8	79.0	1301	13	BX365741	BX365741		C 351	15.4	77.0	600	10	BG800397	BG800397	2102_-29_M	
C 279	15.8	79.0	1432	29	CG356498	P051-4-A0		C 352	15.4	77.0	600	12	BI989303	BI989303	4033_-72_M	
C 280	15.8	79.0	1481	28	CG375444	CH261-191		C 353	15.4	77.0	606	28	AZ633296	AZ633296	1M0486B23	
C 281	15.8	79.0	1560	12	BI413018	602989965		C 354	15.4	77.0	612	29	CE687222	CE687222	l1gr-g8s-	
C 282	15.8	79.0	1674	11	AK053226	Mus_muscu		C 355	15.4	77.0	616	12	BI416184	BI416184	Bj416184	
C 283	15.4	77.0	213	28	BZ096958	CH230-142		C 356	15.4	77.0	618	12	Bu614122	Bu614122	Bu614122	
C 284	15.4	77.0	238	10	BM063490	BM063490		C 357	15.4	77.0	618	14	CK222692	CK222692	204011793	
C 285	15.4	77.0	244	29	CG588373	OST25890		C 358	15.4	77.0	620	14	CB289435	CB289435	V-B-117A0	
C 286	15.4	77.0	261	10	BF554085	UI-R-CO-h		C 359	15.4	77.0	625	14	CA921462	CA921462	EST7639180	
C 287	15.4	77.0	263	12	BU223443	BU223443		C 360	15.4	77.0	629	14	CB036010	CB036010	VVA019P07	
C 288	15.4	77.0	270	10	BA402442	BA402442		C 361	15.4	77.0	631	12	BM494446	BM494446	IPSCBt_1	
C 289	15.4	77.0	273	29	CC596956	CC596956	ZMMBBD039	C 362	15.4	77.0	634	28	CC050144	CC050144	0180533-0	
C 290	15.4	77.0	283	9	AA386834	AA386834	vc22a12.r	C 363	15.4	77.0	636	29	CE628613	CE628613	l1gr-g8s-	
C 291	15.4	77.0	286	12	BQ033999	UI-1-CFO-		C 364	15.4	77.0	639	12	BQ415023	BQ415023	DVIMBKO00	
C 292	15.4	77.0	287	10	BI818148	BI818148		C 365	15.4	77.0	639	29	BX226827	BX226827	Hamo_rer	
C 293	15.4	77.0	300	10	BM264008	BM264008		C 366	15.4	77.0	647	13	B0167199	B0167199	WHR0058_F	
C 294	15.4	77.0	304	10	BM266963	BM266963		C 367	15.4	77.0	648	14	CB059023	CB059023	NISC_jx10	
C 295	15.4	77.0	336	10	BM238391	BM238391		C 368	15.4	77.0	650	11	AK005183	AK005183	Mus_muscu	
C 296	15.4	77.0	351	10	AM470292	xz82b05.x		C 369	15.4	77.0	654	28	CC1892563	CC1892563	02S2027-0	

BI771144	EBrc003	SQ
A2255310	RPCI_-23_-9	
AA396945	mr1f10.x	
AI851749	UI-M-BH0-	
CF506831	USDA-EP_1	
AI216097	gm40b03.x	
B232692	Pa78g04.y	
BY605976	BY605976	
BI766819	ah37a12.s	
AA209714	mw72b02.x	
CF506703	USDA-FP_1	
AI655581	lc27h12.x	
BF061808	7k67C02.x	
BY228504	BY228504	
R49177	Y958a10.g1	
EX189233	Danio_rer	
AA387236	vc21g11.x	
BE944649	UI-M-BH3-	
AA013929	mh24g08.x	
BI113931	BI113931	
CD372378	UI-R-GD0-	
AA881328	vx12h08.x	
AQ0117439	AA0117439	
BY465273	BY465273	
AA013931	mh24h08.x	
BF660500	ma92f12.	
CF156184	CF156184	

370	15.4	77.0	658	29	CG400372	01S0611-1	443	15.2	76.0	263	9	AA506314	TH45C05.8
C 371	15.4	77.0	658	29	CG427792	01S0777-0	C 443	15.2	76.0	264	9	AV238655	AV238655
C 372	15.4	77.0	661	10	BF006633	EST435131	C 444	15.2	76.0	271	9	AV259730	AV259730
C 373	15.4	77.0	663	12	BH592999	BOHUS59TR	C 445	15.2	76.0	277	10	BB010692	BB010692
C 374	15.4	77.0	669	18	BH389950	BH389950	C 446	15.2	76.0	286	10	BB282695	BB282695
C 375	15.4	77.0	671	28	BH765743	BH765743	C 447	15.2	76.0	286	9	AL640424	AL640424
C 376	15.4	77.0	677	9	AA556716	571_Lob10	C 448	15.2	76.0	290	9	AL640424	AL640424
C 377	15.4	77.0	678	9	AA556716	571_Lob10	C 449	15.2	76.0	290	13	BX780974	BX780974
C 378	15.4	77.0	680	10	BB357252	BB357252	C 450	15.2	76.0	293	10	BB005204	BB005204
C 379	15.4	77.0	687	28	BZ495354	BONIT35TF	C 451	15.2	76.0	293	12	BB657032	BB657032
C 380	15.4	77.0	690	14	CF999829	CF999829	C 452	15.2	76.0	295	14	N78343	N78343
C 381	15.4	77.0	695	14	CD003924	AVAN019P07	C 453	15.2	76.0	300	9	AV219173	AV219173
C 382	15.4	77.0	711	28	BH085732	RPCT-24-9	C 454	15.2	76.0	303	9	A1775359	EST256459
C 383	15.4	77.0	712	12	BG523196	29-89_Ste	C 455	15.2	76.0	304	10	BB219213	BB219213
C 384	15.4	77.0	735	28	AQ345909	RPC111-12	C 456	15.2	76.0	310	10	BB205297	BB205297
C 385	15.4	77.0	762	14	CA323241	UI-M-FXO-	C 457	15.2	76.0	311	28	AZ454440	IM0256H20
C 386	15.4	77.0	785	12	BG468157	602509732	C 458	15.2	76.0	318	10	AM902881	OV3-NN102
C 387	15.4	77.0	835	28	BH593136	BOHRQ33TF	C 459	15.2	76.0	318	10	BE839946	RC3-FN014
C 388	15.4	77.0	838	29	CNS02014	AL17657	C 460	15.2	76.0	326	12	BG554416	dab27C07.
C 389	15.4	77.0	840	29	CNS050502	AL132216	C 461	15.2	76.0	331	9	AA650522	nc01611.8
C 390	15.4	77.0	843	29	CNS020902	CH240_368	C 462	15.2	76.0	334	10	BB223035	BB223035
C 391	15.4	77.0	843	29	CNS04041R	AL294264	C 463	15.2	76.0	342	29	CG928415	MBEFH79TF
C 392	15.4	77.0	846	12	BG719954	602691303	C 464	15.2	76.0	344	14	CF506812	USDA-FP_1
C 393	15.4	77.0	848	14	CNS65959	CG1X06823	C 465	15.2	76.0	354	10	BF225297	uz45C11-Y_V
C 394	15.4	77.0	849	29	CNS041C5	ALJ06302	C 466	15.2	76.0	354	29	CEA35568	c1gr-g88-
C 395	15.4	77.0	857	13	BUS58356	AGENCOCURT	C 467	15.2	76.0	358	28	AQ010566	F23HTFB
C 396	15.4	77.0	857	29	CC805534	CG805534	C 468	15.2	76.0	363	13	BE655876	UI-M-BGO-
C 397	15.4	77.0	862	29	CC579888	CH240_374	C 469	15.2	76.0	363	13	BY213050	BE2513050
C 398	15.4	77.0	863	29	CG020988	ZMMBRC055	C 470	15.2	76.0	367	28	BZ592758	SALK_0291
C 399	15.4	77.0	864	14	CF829949	UCRCRO1_0	C 471	15.2	76.0	367	10	AM447338	88787_MAR
C 400	15.4	77.0	870	29	CGS17833	CH240_364	C 472	15.2	76.0	374	14	CD195151	MS1-00887
C 401	15.4	77.0	873	29	CG954789	MBBDH02TR	C 473	15.2	76.0	377	13	BY166002	BY166002
C 402	15.4	77.0	889	12	BG300625	HVSMED001	C 474	15.2	76.0	379	10	AM490991	UI-M-BH3-
C 403	15.4	77.0	887	13	BX845255	BX845255	C 475	15.2	76.0	380	28	AQ049765	CGR-64A3-
C 404	15.4	77.0	896	14	CG560879	AGENCOCURT	C 476	15.2	76.0	387	10	AM542363	AM542363
C 405	15.4	77.0	912	29	CG560879	AGENCOCURT	C 477	15.2	76.0	390	13	BY444993	BY444993
C 406	15.4	77.0	922	13	BQ737035	AGENCOCURT	C 478	15.2	76.0	392	10	AM697635	ST64G03_P
C 407	15.4	77.0	922	13	BY746278	BY746278	C 479	15.2	76.0	395	28	AQ637841	RPCT-11-4
C 408	15.4	77.0	952	13	BX341013	BX341013	C 480	15.2	76.0	399	28	AQ673057	AQ673057
C 409	15.4	77.0	980	29	CNS022078	AL21388	C 481	15.2	76.0	404	14	CB811908	AMGNNUCIN
C 410	15.4	77.0	986	29	CG900580	ZMMBRC050	C 482	15.2	76.0	407	10	BF548993	BF548993
C 411	15.4	77.0	1031	12	B1834473	603084638	C 483	15.2	76.0	410	10	AM343233	AM343233
C 412	15.4	77.0	1201	9	AL513771	AL513771	C 484	15.2	76.0	411	13	BY006359	BY006359
C 413	15.4	77.0	1201	9	AL564116	AL564116	C 485	15.2	76.0	414	14	CD012333	CD012333
C 414	15.4	77.0	1348	12	BM047112	603627450	C 486	15.2	76.0	420	28	AQ205240	AQ205240
C 415	15.4	77.0	1548	11	AK0477338	Mus muscu	C 487	15.2	76.0	421	14	CB811908	CB811908
C 416	15.4	77.0	1514	28	CC291468	CC291468	C 488	15.2	76.0	422	14	CF372888	CF372888
C 417	15.4	77.0	1514	28	CC291468	CH261-173	C 489	15.2	76.0	422	14	T90278	T90278
C 418	15.4	77.0	154	11	AK089011	Mus muscu	C 490	15.2	76.0	429	10	AM484023	AM484023
C 419	15.4	77.0	154	11	AK089011	Mus muscu	C 491	15.2	76.0	430	29	CC291468	CC291468
C 420	15.4	77.0	154	11	AK044408	Mus muscu	C 492	15.2	76.0	431	10	AM479605	AM479605
C 421	15.4	77.0	154	11	AK054548	Mus muscu	C 493	15.2	76.0	431	10	BB848002	BB848002
C 422	15.4	77.0	154	11	AK047026	Mus muscu	C 494	15.2	76.0	434	14	CD195134	CD195134
C 423	15.4	77.0	154	11	AK050673	Mus muscu	C 495	15.2	76.0	436	14	BB449928	BB449928
C 424	15.4	77.0	154	11	AK036637	Mus muscu	C 496	15.2	76.0	436	14	CB230886	CB230886
C 425	15.4	77.0	154	11	AK034000	Mus muscu	C 497	15.2	76.0	438	9	A1601572	A1601572
C 426	15.2	76.0	127	12	B1524648	603051625	C 498	15.2	76.0	439	10	BE626665	BE626665
C 427	15.2	76.0	136	28	BZ613777	KB-RH143C0	C 499	15.2	76.0	443	9	A1580295	A1580295
C 428	15.2	76.0	163	29	CE605068	CE605068	C 500	15.2	76.0	443	9	A1601294	A1601294
C 429	15.2	76.0	193	28	AA227421	1M0050A16	C 501	15.2	76.0	446	10	BB726107	BB726107
C 430	15.2	76.0	212	10	BB133381	BB133381	C 502	15.2	76.0	451	12	BJ387849	BJ387849
C 431	15.2	76.0	225	29	CE272099	c1gr-g88-	C 503	15.2	76.0	451	9	A1022508	A1022508
C 432	15.2	76.0	230	9	AL888947	AL888947	C 504	15.2	76.0	451	28	AQ668567	HS_5331_B
C 433	15.2	76.0	230	9	AV274759	AV274759	C 505	15.2	76.0	451	28	BF559336	UI-R-E0-C
C 434	15.2	76.0	230	9	AV529468	AV529468	C 506	15.2	76.0	454	10	AM984583	AM984583
C 435	15.2	76.0	231	10	BB216715	BB216715	C 507	15.2	76.0	456	12	BG816750	dab67F02.
C 436	15.2	76.0	233	10	BB219461	BB219461	C 508	15.2	76.0	457	9	AA917696	AA917696
C 437	15.2	76.0	243	13	BX779126	BX779126	C 509	15.2	76.0	458	10	AM648757	AM648757
C 438	15.2	76.0	244	9	AV299150	AV299150	C 510	15.2	76.0	460	28	AQ080159	HS_3181_A
C 439	15.2	76.0	246	9	AV293556	AV293556	C 511	15.2	76.0	462	13	BUS78131	ga147e05
C 440	15.2	76.0	259	12	BM118138	BM118138	C 512	15.2	76.0	463	28	BZ825560	BZ825560
C 441	15.2	76.0	261	9	AV352157	AV352157	C 513	15.2	76.0	464	9	AUT175638	AUT175638
C 442	15.2	76.0	261	13	B0896780	X045F03_P	C 514	15.2	76.0	469	9	AL890531	AL890531
							C 515	15.2	76.0	471	12	BM117916	BM117916



516	15.2	76.0	472	9	A1189543	A1189543 qd22h02.x	C 589	15.2	76.0	574	10	BE354681	BE354681 EST354771
517	15.2	76.0	472	9	A1838722	A1838722 UT-M-A00-	C 590	15.2	76.0	575	28	A2838089	A2838089 2M0133G08
518	15.2	76.0	472	28	A0660554	A0660554 Sheared D	C 591	15.2	76.0	576	12	BG885427	BG885427 daad2c08
519	15.2	76.0	474	10	BE623205	BE623205 ut95a06.x	C 592	15.2	76.0	577	10	AM138767	AM138767 UT-H-B11-
520	15.2	76.0	475	14	CB005274	CB005274 VVC01B03	C 593	15.2	76.0	578	12	B0359158	B0359158 BJ359158
521	15.2	76.0	477	9	A1935668	A1935668 wo99c07.x	C 594	15.2	76.0	579	12	B0359316	B0359316 BJ359316
522	15.2	76.0	478	12	BI322592	BI322592 kx13h07.y	C 595	15.2	76.0	579	29	CE627130	CE627130 L1gr-g88-
523	15.2	76.0	478	13	BK635970	BK635970 BX635970	C 596	15.2	76.0	580	13	CE627105	CE627105 SJM2BJB08
524	15.2	76.0	479	10	BE114484	BE114484 UT-R-CA0-	C 597	15.2	76.0	581	14	CB916619	CB916619 VVD111G05
525	15.2	76.0	481	28	BE2290	BE2290 RRC11-14H1	C 598	15.2	76.0	583	10	BE694232	BE694232 602082857
526	15.2	76.0	489	9	AA859269	AA859269 UT-R-E0-C	C 599	15.2	76.0	586	12	B0361629	B0361629 BJ361629
527	15.2	76.0	494	9	A1531862	A1531862 SD03145.5	C 600	15.2	76.0	586	29	CC466857	CC466857 CH240_136
528	15.2	76.0	496	13	BO556161	BO556161 H4038D11-	C 601	15.2	76.0	587	12	CB425375	CB425375 600352 MA
529	15.2	76.0	497	28	AOB33959	AOB33959 HS 5261 A	C 602	15.2	76.0	588	14	CA892704	CA892704 B0174B03-
530	15.2	76.0	498	12	BJ388628	BJ388628 BJ388628	C 603	15.2	76.0	591	10	CA367755	CA367755 60128547
531	15.2	76.0	499	14	W20132	W20132 2B40C10.r1	C 604	15.2	76.0	591	14	CA367755	CA367755 643750 NC
532	15.2	76.0	504	12	BJ363161	BJ363161 BJ363161	C 605	15.2	76.0	591	14	CA404168	CA404168 EL01N0513
533	15.2	76.0	504	14	CA208521	CA208521 SCACSB111	C 606	15.2	76.0	591	14	CB425375	CB425375 600352 MA
534	15.2	76.0	506	28	AQ243748	AQ243748 HS 2057 B	C 607	15.2	76.0	592	12	BJ389256	BJ389256 BJ389256
535	15.2	76.0	507	9	A1657550	A1657550 fcl5a02.y	C 608	15.2	76.0	593	29	CC743217	CC743217 ZMMBBD011
536	15.2	76.0	509	9	AA831831	AA831831 oc85g03.B	C 609	15.2	76.0	594	14	CB910032	CB910032 VVD173B01
537	15.2	76.0	509	28	AQ776625	AQ776625 HS_2148 B	C 610	15.2	76.0	595	14	CA387084	CA387084 668874 NC
538	15.2	76.0	510	14	CB424865	CB424865 599499 MA	C 611	15.2	76.0	596	13	B0357680	B0357680 603476801
539	15.2	76.0	510	14	CF403237	CF403237 CSECS003F	C 612	15.2	76.0	597	29	CE795949	CE795949 L1gr-g88-
540	15.2	76.0	511	12	BJ363045	BJ363045 BJ363045	C 613	15.2	76.0	600	28	BE708259	BE708259 OGBAP06TC
541	15.2	76.0	519	28	AO505167	AO505167 RPCI-11-2	C 614	15.2	76.0	601	9	AT062238	AT062238 GH01378_5
542	15.2	76.0	520	13	BU080919	BU080919 UL54PH12	C 615	15.2	76.0	601	29	DF463195	DF463195 Datio_rer
543	15.2	76.0	520	14	CA545244	CA545244 K0110E05-	C 616	15.2	76.0	602	14	CR486121	CR486121 POL1_35-G
544	15.2	76.0	520	14	CB715589	CB715589 AMGNNUC:N	C 617	15.2	76.0	607	12	BG467408	BG467408 EST12 Pac
545	15.2	76.0	523	28	BH117649	BH117649 RPCI-24-3	C 618	15.2	76.0	607	14	CD646919	CD646919 AUT_106 I
546	15.2	76.0	524	12	BM256385	BM256385 518740 MA	C 619	15.2	76.0	609	9	AI062055	AI062055 GH01103_5
547	15.2	76.0	525	14	CB720171	CB720171 AMGNNUC-U	C 620	15.2	76.0	609	9	AJ397505	AJ397505 AJ397505
548	15.2	76.0	525	28	AQ940449	AQ940449 Sheared D	C 621	15.2	76.0	609	13	B0818793	B0818793 UA47BEP05
549	15.2	76.0	526	28	CC064938	CC064938 Igm04004D	C 622	15.2	76.0	609	13	BX267334	BX267334 BA267334
550	15.2	76.0	529	10	AM690744	AM690744 NF033A07S	C 623	15.2	76.0	610	28	AQ955222	AQ955222 LERAB06TF
551	15.2	76.0	530	18	AO554837	AO554837 RPCI-11-3	C 624	15.2	76.0	615	9	AM034396	AM034396 EST77967
552	15.2	76.0	534	14	CA545539	CA545539 K0115F12-	C 625	15.2	76.0	615	12	BG791233	BG791233 ESTFN1096
553	15.2	76.0	535	29	CG925610	CG925610 MBENJ23TR	C 626	15.2	76.0	617	13	BX481654	BX481654 DKFED686C
554	15.2	76.0	536	14	CA741071	CA741071 wem1C.pko	C 627	15.2	76.0	617	28	A2412192	A2412192 1M0185K01
555	15.2	76.0	537	10	BE749525	BE749525 200496 MA	C 628	15.2	76.0	619	28	A2423496	A2423496 1M0202C09
556	15.2	76.0	540	28	BH015619	BH015619 TDGCC86TH	C 629	15.2	76.0	621	12	BG577046	BG577046 602599461
557	15.2	76.0	542	10	AM237816	AM237816 xm81h09.x	C 630	15.2	76.0	621	14	CA394142	CA394142 c847905.y
558	15.2	76.0	542	10	BE015099	BE015099 127277 MA	C 631	15.2	76.0	621	28	A2298192	A2298192 RPCI-23-1
559	15.2	76.0	546	12	BI501358	BI501358 kx31c03.y	C 632	15.2	76.0	625	13	BU359964	BU359964 603476969
560	15.2	76.0	546	12	BM068107	BM068107 KS08017AL	C 633	15.2	76.0	627	12	B0325128	B0325128 BJ325128
561	15.2	76.0	547	10	AM792048	AM792048 D00974-R	C 634	15.2	76.0	629	28	AO642252	AO642252 RPCI93-IDP
562	15.2	76.0	548	9	A1813671	A1813671 wJ65d10.x	C 635	15.2	76.0	630	10	AM646759	AM646759 EST327213
563	15.2	76.0	549	14	CD854887	CD854887 DHOAMM6Z	C 636	15.2	76.0	630	29	CE496250	CE496250 L1gr-g88-
564	15.2	76.0	549	28	AO992197	AO992197 nbeD0085A	C 637	15.2	76.0	632	12	BG378240	BG378240 UT-R-CV1-
565	15.2	76.0	550	9	A1934581	A1934581 wp09g01.x	C 638	15.2	76.0	633	13	BU402588	BU402588 604140873
566	15.2	76.0	550	12	BI131118	BI131118 G115P631	C 639	15.2	76.0	633	28	B2372491	B2372491 1e68B04.D
567	15.2	76.0	550	14	CD647519	CD647519 AUF_108 D	C 640	15.2	76.0	634	28	B2881424	B2881424 CH240_266
568	15.2	76.0	551	12	BI186293	BI186293 BFL26_000	C 641	15.2	76.0	638	13	BX267333	BX267333 BX267333
569	15.2	76.0	551	12	BU359641	BU359641 BJ359641	C 642	15.2	76.0	638	14	CB321897	CB321897 UT-CF-FNO
570	15.2	76.0	553	10	AME57940	AME57940 L0288B11-	C 643	15.2	76.0	639	28	BH633624	BH633624 SALK_0428
571	15.2	76.0	554	14	CP162722	CP162722 B0715H02-	C 644	15.2	76.0	640	10	BB021763	BB021763 BR021763
572	15.2	76.0	555	14	CB613545	CB613545 AMGNNUC-U	C 645	15.2	76.0	640	14	CD190714	CD190714 WS1_0064U
573	15.2	76.0	555	28	BZ708270	BZ708270 OGBAP06TM	C 646	15.2	76.0	640	28	B2925525	B2925525 CH240_72E
574	15.2	76.0	557	29	CG971340	CG971340 MBERG08TF	C 647	15.2	76.0	643	12	B1386594	B1386594 BFL26_001
575	15.2	76.0	560	13	BU948965	BU948965 ln71e03.y	C 648	15.2	76.0	643	12	B1649022	B1649022 603277202
576	15.2	76.0	561	28	AZ054783	AZ054783 RPCI-23-4	C 649	15.2	76.0	646	28	B2805871	B2805871 PUGAL54TB
577	15.2	76.0	563	9	A1658278	A1658278 f620h09.y	C 650	15.2	76.0	647	28	AZ804958	AZ804958 2M0066D12
578	15.2	76.0	564	10	AM147848	AM147848 dai0e06.x	C 651	15.2	76.0	647	29	BX134945	BX134945 Datio_rer
579	15.2	76.0	564	14	CD731736	CD731736 4041736 I	C 652	15.2	76.0	650	13	BX667941	BX667941 BX667941
580	15.2	76.0	566	12	BI964609	BI964609 1e63g04.y	C 653	15.2	76.0	650	28	B2448956	B2448956 BONGM57TR
581	15.2	76.0	566	14	CB098319	CB098319 kx52b01.y	C 654	15.2	76.0	652	10	BE056531	BE056531 894010A05
582	15.2	76.0	569	10	BE475053	BE475053 sp71f10.y	C 655	15.2	76.0	653	14	CA501293	CA501293 WHR4031_G
583	15.2	76.0	569	13	BM291870	BM291870 BM291970	C 656	15.2	76.0	656	10	BE670697	BE670697 7e38h07_x
584	15.2	76.0	569	29	CB828021	CB828021 L1gr-g88-	C 657	15.2	76.0	658	13	BY738487	BY738487 BY738487
585	15.2	76.0	571	14	CA825254	CA825254 R56A08 tw	C 658	15.2	76.0	659	10	AM612120	AM612120 h994d08.x
586	15.2	76.0	572	29	CE694107	CE694107 L1gr-g88-	C 659	15.2	76.0	660	13	BQ240611	BQ240611 TAB805015B
587	15.2	76.0	573	10	AM474485	AM474485 xs97d01.x	C 660	15.2	76.0	661	29	CG219902	CG219902 OG4AL10TC
588	15.2	76.0	573	29	AG331466	AG331466 Lotus cor	C 661	15.2	76.0	662	14	CD719814	CD719814 VVB165D11



LOCUS BG201597 219 bp mRNA linear EST 21-APR-2001  
 DEFINITION RST0938 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG201597  
 VERSION BG201597.1 GI:137323284  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 219)  
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.  
 TITLE Creation of genome-wide protein expression libraries using random activation of gene expression  
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
 MEDLINE 21227151  
 PUBMED 11329013  
 COMMENT Contact: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@athersys.com  
 High quality sequence stop: 219.  
 LOCATION/Qualifiers  
 1..219  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="HT1080"  
 /clone\_lib="Athersys RAGE Library"  
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN  
 Query Match 100.0%; Score 20; DB 12; Length 219;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
 |||  
 72 GGTCCTACTCATGTTGATG 53

Db 72 GGTCCTACTCATGTTGATG 53

RESULT 2  
 LOCUS BG212172 221 bp mRNA linear EST 21-APR-2001  
 DEFINITION RST1643 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG212172  
 VERSION BG212172.1 GI:13733747  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 221)  
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.  
 TITLE Creation of genome-wide protein expression libraries using random activation of gene expression  
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
 MEDLINE 21227151  
 PUBMED 11329013

COMMENT Contact: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@athersys.com  
 High quality sequence stop: 221.  
 LOCATION/Qualifiers  
 1..221  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="HT1080"  
 /clone\_lib="Athersys RAGE Library"  
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN  
 Query Match 100.0%; Score 20; DB 12; Length 221;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
 |||  
 74 GGTCCTACTCATGTTGATG 55

Db 74 GGTCCTACTCATGTTGATG 55

RESULT 3  
 LOCUS BG185555 352 bp mRNA linear EST 21-APR-2001  
 DEFINITION RST4504 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG185555  
 VERSION BG185555.1 GI:13707242  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 352)  
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramchandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.  
 TITLE Creation of genome-wide protein expression libraries using random activation of gene expression  
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
 MEDLINE 21227151  
 PUBMED 11329013  
 COMMENT Contact: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@athersys.com  
 High quality sequence stop: 352.  
 LOCATION/Qualifiers  
 1..352  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="HT1080"  
 /clone\_lib="Athersys RAGE Library"  
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 352;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
|||||  
DB 283 GGTCCTACTCATGTTGATG 302

RESULT 4  
CB852193/c 359 bp mRNA linear EST 22-APR-2003  
LOCUS  
DEFINITION UI-CF-FNO-seq-1-02-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone  
UI-CF-FNO-seq-1-02-0-UI 3', mRNA sequence.  
ACCESSION CB852193  
VERSION CB852193.1 GI:30047109  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 359)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 249-275, >AT rich/low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLY(A=NO.

## FEATURES

source  
1..359  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-FNO-seq-1-02-0-UI"  
/issue\_type="Human Lung Epithelial cells"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_1ib="UI-CF-FNO"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-CF-FNO is a subtracted cDNA library derived from two  
normalized Human Lung epithelial cell libraries (EN1 and  
DUI) The library was subtracted according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. For additional information, contact:  
bento-soares@uiowa.edu  
TAG\_SEQ=None found"

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 359;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCCTACTCATGTTGATG 20  
|||||

DB 149 GGTCCTACTCATGTTGATG 130

RESULT 5  
BM796479/c 370 bp mRNA linear EST 05-MAR-2002  
LOCUS  
DEFINITION K-EST0079212 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-70-A04  
5', mRNA sequence.  
ACCESSION BM796479  
VERSION BM796479.1 GI:19144711  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 370)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontler Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.krribb.re.kr  
Plate: 70 row: A column: 04  
High quality sequence stop: 370.  
Location/Qualifiers  
1..370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S22SNU16n1-70-A04"  
/sex="F"  
/tissue\_type="Ascites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-16"  
/lab\_host="DH10B"  
/clone\_1ib="S22SNU16n1"  
/note="Organ: Stomach; Vector: pT73-Pac; Site 1: EcoRI;  
Site 2: NotI; The S22SNU16 library was contributed by the  
Soares laboratory and it was constructed as described by  
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
Research 6(9): 791-806. RNA was prepared from harvested  
cells of SNU-16 culture. SNU-16 cell was obtained from  
Korean Cell Line Bank (KCLB). SNU-16 was established from  
ascitic fluids of Korean patients by Park J.G. et al.  
(1990), Cancer Res 50: 2773-2780."

## FEATURES

source  
1..370  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S22SNU16n1-70-A04"  
/sex="F"  
/tissue\_type="Ascites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-16"  
/lab\_host="DH10B"  
/clone\_1ib="S22SNU16n1"  
/note="Organ: Stomach; Vector: pT73-Pac; Site 1: EcoRI;  
Site 2: NotI; The S22SNU16 library was contributed by the  
Soares laboratory and it was constructed as described by  
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
Research 6(9): 791-806. RNA was prepared from harvested  
cells of SNU-16 culture. SNU-16 cell was obtained from  
Korean Cell Line Bank (KCLB). SNU-16 was established from  
ascitic fluids of Korean patients by Park J.G. et al.  
(1990), Cancer Res 50: 2773-2780."

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 370;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
|||||  
DB 121 GGTCCTACTCATGTTGATG 102

RESULT 6  
AV687532/c 408 bp mRNA linear EST 16-JAN-2002  
LOCUS  
DEFINITION AV687532 GKC Homo sapiens cDNA clone GKDC01 5', mRNA sequence.  
ACCESSION AV687532  
VERSION AV687532.1 GI:10289395  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 408)  
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Guo,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z., and Han,Z.  
TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
MEDLINE 21625106  
PUBMED 11752456  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1..408  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GKDCDA01"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_1ib="GKC"  
/note="Vector: pBlueScript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 408;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
120 GGTCCATCTCATGTTGATG 101

Db

RESULT 7  
AV690967/c 408 bp mRNA linear EST 16-JAN-2002  
DEFINITION AV690967 GKC Homo sapiens cDNA clone GKDCDD05 5', mRNA sequence.  
ACCESSION AV690967  
VERSION AV690967.1 GI:10292830  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 408)  
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Guo,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,W., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z., and Han,Z.  
TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
MEDLINE 21625106  
PUBMED 11752456  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
Location/Qualifiers

source  
1..408  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GKDCDD05"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_1ib="GKC"  
/note="Vector: pBlueScript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 408;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
120 GGTCCATCTCATGTTGATG 101

Db

RESULT 8  
A1961077 437 bp mRNA linear EST 08-MAR-2000  
LOCUS A1961077  
DEFINITION w444901.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2474160 3' similar to gb:D14533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS (HUMAN);, mRNA sequence.  
ACCESSION A1961077  
VERSION A1961077.1 GI:5753858  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 437)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@db-ir@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: www-bio.ln1.gov/bbtp/image/image.html  
Insert Length: 766 Std Error: 0.00  
Seq primer: -40up from Gibco.

FEATURES  
source  
1..437  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2474160"  
/tissue\_type="pooled germ cell tumors"  
/lab\_host="FDH10B"  
/clone\_1ib="NCI CGAP GC6"  
/note="Vector: p7773D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP GC4 was prepared, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 1257096-1258631, 1469064-1470963, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo.

ORIGIN  
Location/Qualifiers

Query Match		100.0%	Score 20;	DB 9;	Length 437;
Best Local Similarity		100.0%;	Pred. No. 75;		
Matches	20;	Conservative	0;	Mismatches	0; Gaps 0;
OY	1	GATCCATCTCATGTTGATG	20		
Db	158	GATCCATCTCATGTTGATG	177		
RESULT 9					
BUE679004					
LOCUS	BUE679004	461 bp.	mRNA	linear	EST 07-OCT-2002
DEFINITION	UI-CF-DUI-aac-o-05-0-ui.s1 UI-CF-DUI Homo sapiens cDNA clone				
ACCESSION	BUE679004				
VERSION	BUE679004.1	GI:23526515			
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 461) Bonaldo,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996) JOURNAL MEDLINE PUBMED Contact: McCray, PB 8889548				

University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics ([www.reegen.com](http://www.reegen.com)) or from Open Biosystems  
([www.openbiosystems.com](http://www.openbiosystems.com)).  
The following repetitive elements were found in this cDNA  
sequence: 49-75, >AT rich#low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA=yes.

FEATURES	Location/Qualifiers
source	1. .461

```

1. .461
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-aat-o-05-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DI108 (Life Technologies) (T1 phage resistant)"
/clone_id="UI-CF-DUI"
/notes="Organ. Lung. Vector: pT73-Pac (Pharmacia) with a
modified polylinker. Site_1: EcoR I; Site_2: Not I,
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT73-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (drl) tail. The sequence tag for this
library is GCGTCAGGC.
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_LIB=UI-CF-DUI
TAG_SEQ=GCGTCAGGC"

```

```

ORIGIN
Query Match          100.0%;   Score 20;   Length 461;
Best Local Similarity 100.0%;   Pred. No. 77;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GGTCCTACTCATCTGTGATG 20
        |||||
Db       175 GGTCCTACTCATCTGTGATG 194

```

RESULT 10	
AA828925	
LOCUS	AA828925 486 bp mRNA linear EST 07-APR-1998
DEFINITION	cd0a10.0.1 NCI CGAP Ov-2 Homo sapiens cDNA clone IMAGE:1374330 similar to gb:D14533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS (HUMAN), mRNA sequence.
ACCESSION	AA828925
VERSION	AA828925.1 GI:2902024
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 486)
REFERENCE	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
AUTHORS	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE	Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Kitzman, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E.B. Consortium/ILNIB at: [www-bio.llnl.gov/bdbp/image/image.html](http://www-bio.llnl.gov/bdbp/image/image.html)  
Insert Length: 528 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 353.

FEATURES	Location/Qualifiers
source	1. .486

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1374330"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/clone_id="NCI CGAP Ovt2"
/note="Vector: pAMP10; mRNA made from invasive ovarian tumor, CDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

```

ORIGIN		Query Match	Score 20;	DB 9;	Length 486;
Beet	Local Similarity	100.0%;	Pred. No. 78;		
Matches	20; Conservative	0;	Mismatches	0;	Indels 0;
Qy	1 GGTCCATCTCATGTTGATG	20			
Db	149 GGTCCATCTCATGTTGATG	168			

RESULT 11	LOCUS	DEFINITION
BI793276/c	487 bp	mRNA linear
BI793276	487 bp	mRNA linear
1650b09.y1	487 bp	mRNA linear
CDNA clone IMAGE:5670113	487 bp	mRNA linear
5'	487 bp	mRNA linear
similar to SW:XPB_HUMAN	487 bp	mRNA linear
P23025	487 bp	mRNA linear
EST 12-MAR-2001	487 bp	mRNA linear
1	487 bp	mRNA linear
Homo sapiens	487 bp	mRNA linear

DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS ;, mRNA sequence.  
 ACCESSION B1793276  
 VERSION B1793276.1 GI:15821001  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 487)  
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Breselli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A., Schmitt,A., Theising,B., Rilter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarelshvili,R., Williams,T., Jackson,Y. and Bowers,Y.  
 TITLE Endocrine Pancreas Consortium  
 JOURNAL Unpublished (2000)  
 COMMENT Other ESTs: 165D09.x1  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu) This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 261.  
 Location/Qualifiers  
 1..487  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5670113"  
 /sex="Both"  
 /tissue\_type="Islets of Langerhans"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B"  
 /clone.lib="Melton Normalized Human Islet 4 N4-HIS 1"  
 /note="Organ: Pancreas; Vector: pSPORT1; Site\_1: Not 1; Site\_2: Sal I; Starting library constructed using Superscript Plasmid library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."  
 ORIGIN  
 Query Match 100.0%; Score 20; DB 12; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 78;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGTCCATCTCATGTTGATG 20  
 |||||  
 DB 459 GGTCCATCTCATGTTGATG 440  
 RESULT 12  
 LOCUS AA167098/c 506 bp mRNA linear EST 09-MAR-1998  
 DEFINITION zp05g10.r1 StrataGene ovarian cancer (#937219) Homo sapiens cDNA

clone IMAGE:595554 5' similar to gb:D14533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS (HUMAN);, mRNA sequence.  
 ACCESSION AA167098  
 VERSION AA167098.1 GI:1745474  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 506)  
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Joet,S., Kirizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Stepcoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
 TITLE WashU-NCI human EST Project  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 667 Std Error: 0.00  
 Seq primer: -28M13 rev2 from Amersham  
 High quality sequence stop: 364.  
 Location/Qualifiers  
 1..506  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:595554"  
 /sex="Female"  
 /dev\_stage="adult, 64 years"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /clone.lib="Stratagene ovarian cancer (#937219)"  
 /note="Vector: Bluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Papillary serous carcinoma, isolated from ascites, 64 year old caucasian. Average insert size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"  
 ORIGIN  
 Query Match 100.0%; Score 20; DB 9; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGTCCATCTCATGTTGATG 20  
 |||||  
 DB 246 GGTCCATCTCATGTTGATG 227  
 RESULT 13  
 LOCUS BQ186624/c 510 bp mRNA linear EST 30-APR-2002  
 DEFINITION UI-E-EJ1-ej-r-n-03-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone  
 ACCESSION BQ186624  
 VERSION BQ186624.1 GI:20362175  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 510)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548

COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.research.com).  
The following repetitive elements were found in this cDNA sequence: 432-458, >AT rich#Low\_complexity (matched complement)  
Seq primer: M13 REVERSE.

FEATURES  
source  
1..510  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ1-ajr-n-03-0-UI"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-EJ1"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-E-EJ1 is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGATCAGAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN  
Query Match 100.0%; Score 20; DB 13; Length 510;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGTCATCTCATGTTGATG 20  
|||||  
Db 332 GGTCCTACTCATGTTGATG 313

RESULT 14  
AA166911 513 bp mRNA linear EST 09-MAR-1998  
LOCUS AA166911  
DEFINITION zpo5g10.s1 StrataGene ovarian cancer (#937219) Homo sapiens cDNA  
clone IMAGE:595554 3 similar to gb:D14533 DNA-REPAIR PROTEIN  
COMPLEMENTING XP-A CELLS (HUMAN); contains MER37.ct2 MER37 repetitive  
element ; mRNA sequence.  
AA166911  
VERSION AA166911.1 GI:1745421  
EST.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 513)  
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Kizman,D., Kucaba,T., Lacy,W., Le,N., Lennon,G., Marra,M.,

TITLE  
JOURNAL  
COMMENT  
Martin,J., Moore,B., Schejlenberg,K., Steptoe,M., Tan,F.,  
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
WashU-NCI human EST Project  
Unpublished (1997)  
Contact: Wilson RK  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 667 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 263.

FEATURES  
source  
1..513  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:595554"  
/sex="female"  
/dev\_stage="adult, 64 years"  
/lab\_host="SOBR (Kanamycin resistant)"  
/clone\_lib="Stratagene ovarian cancer (#937219)"  
/note="Vector: Bluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dT.  
Papillary serous carcinoma, isolated from ascites, 64 year  
old caucasian. Average insert size: 0.8 kb; Uni-ZAP XR  
Vector; -5' adaptor sequence: 5' GAATTCGACGACGAG 3' -3'  
adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 513;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGTCATCTCATGTTGATG 20  
|||||  
Db 357 GGTCATCTCATGTTGATG 376

RESULT 15  
B0184648 515 bp mRNA linear EST 30-APR-2002  
LOCUS B0184648  
DEFINITION UI-E-EJ1-ajr-n-03-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone  
UI-E-EJ1-ajr-n-03-0-UI 3', mRNA sequence.  
B0184648  
VERSION B0184648.1 GI:20360199  
EST.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 515)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)

JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research



Genetics (www.reegen.com).  
The following repetitive elements were found in this cDNA  
sequence: 49-75, >AT rich#low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA=yes

#### FEATURES

source

Location/Qualifiers  
1..515  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-B-EJ1-ajr-n-03-0-UI"  
/tissue\_type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina Foveal and Macular, RPE and  
Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-B-EJ1"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-B-EJ1 is a subtracted cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT7T3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are: fetal eyes,  
AGATTCAGAGA; lens, CGATTAGCGA; eye anterior segment,  
AATGCCGAT; optic nerve, CCATTAGTG; retina, CCCG; Retina  
Foveal and Macular, GTCG; RPE and Choroid, ACCTA. This  
library was created for the program, Gene Discovery in the  
visual system, supported by National Eye Institute (NEI).  
TAG\_LIB=UI-B-EJ1  
TAG\_SEQ=GTCG"

#### ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 515;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTGATG 20  
|||||  
Db 175 GGTCCATCTCATGTGTGATG 194

RESULT 16  
AM890854 534 bp mRNA linear EST 24-MAY-2000  
DEFINITION RC5-NT0053-100500-021-G11 NT0053 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AM890854  
VERSION AM890854.1 GI:8055059  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 534)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zagzo, M.A., Bordin, S., Costa, F.P.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

#### TITLE

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202653

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/ICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC5-NT0053-100  
500-021-G1453=2000-05-10x4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 12  
High quality sequence stop: 486.

#### FEATURES

source

Location/Qualifiers  
1..534  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="NT0053"  
/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORSTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

#### ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 534;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTGTGATG 20  
|||||  
Db 430 GGTCCATCTCATGTGTGATG 449

RESULT 17  
CB048344 587 bp mRNA linear EST 17-JAN-2003  
LOCUS CB048344  
DEFINITION NISC\_G104C05\_Y1 NCI\_CGAP\_P128 Homo sapiens cDNA clone IMAGE:3270537  
5', mRNA sequence.  
ACCESSION CB048344  
VERSION CB048344.1 GI:27786631  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 587)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov

#### COMMENT

cDNA library preparation:  
DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL  
Sequencing Center (NISC)  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
info@image.llnl.gov  
Plate: LLM8006 row: F column: 10  
Seq primer: M13Rpl reverse primer (ABI).

#### FEATURES

source

Location/Qualifiers  
1..587  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3270537"  
/sex="male"

/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr28"  
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)  
with a modified polylinker; Plasmid DNA from the  
normalized library NCI CGAP Pr22 was prepared, and as  
circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from a pool  
of 5,000 clones made from the same library (clonoids  
985608-986759, 110192-110199, and 1217928-1220615).  
Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 587;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
Db 214 GGTCCATCTCATGTTGATG 195

RESULT 18  
CB048343 591 bp mRNA linear EST 17-JUN-2003  
LOCUS NISC g104c05.x1 NCI CGAP Pr28 Homo sapiens CDNA IMAGE:3270537  
DEFINITION 3', mRNA sequence.  
ACCESSION CB048343  
VERSION CB048343.1 GI:27786630  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 591)  
AUTHORS NCI-CGAP htcp://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgep@bbs-r@mail.nih.gov  
CDNA Library Preparation:  
DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL  
Sequencing Center (NISC)  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
info@image.llnl.gov  
Plate: LLA8006 row: F column: 10  
Seq primer: -21m3 forward primer (ABT).  
Location/Qualifiers

FEATURES  
source 1..591

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3270537"  
/sex="male"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr28"  
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)  
with a modified polylinker; Plasmid DNA from the  
normalized library NCI CGAP Pr22 was prepared, and as  
circles were made in vitro. Following HAP purification,  
this DNA was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from a pool  
of 5,000 clones made from the same library (clonoids  
985608-986759, 110192-110199, and 1217928-1220615).  
Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 591;

Best Local Similarity 100.0%; Pred. No. 85;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
Db 378 GGTCCATCTCATGTTGATG 397

RESULT 19  
AM979099 592 bp mRNA linear EST 02-JUN-2000  
LOCUS EST9391209 MAGE resequences, MAGP Homo sapiens CDNA, mRNA sequence.  
DEFINITION AM979099  
ACCESSION AM979099  
VERSION AM979099.1 GI:8170384  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 592)  
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and  
Quackenbush, J.

TITLE Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
JOURNAL Unpublished (2000)  
COMMENT Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: john@igf.org  
Plate: 402  
Seq primer: Forward.  
Location/Qualifiers

FEATURES  
source 1..592

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGP"  
/note="Vector: pBluescriptsm"

## ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 592;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCCATCTCATGTTGATG 20  
|||||  
Db 270 GGTCCATCTCATGTTGATG 289

RESULT 20  
BM739320 595 bp mRNA linear EST 01-MAR-2002  
LOCUS K-EST0008760 S2SNUG68 Homo sapiens CDNA clone S2SNUG68-6-H08 5',  
DEFINITION mRNA sequence.  
ACCESSION BM739320  
VERSION BM739320.1 GI:19060649  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 595)  
AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.W., Park, H.S., Kim, S. and  
Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 6 row: H column: 08  
High quality sequence stop: 595.  
Location/Qualifiers

# FEATURES

source

```
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SSNU668-6-H08"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-668"
/lab_host="Top10F"
/clone_11b="SSNU668"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The cDNA vector
vector was adjusted to have about 60nt. The dt-tailed
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F by electroporation
method."
```

# ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 595;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCCATCTCATGTGTGATG 20  
|||||  
Db 412 GGTCCATCTCATGTGTGATG 393

RESULT 21  
B1792983 598 bp mRNA linear EST 12-MAR-2002  
LOCUS i650b09.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens  
DEFINITION cDNA clone IMAGE:5670113 similar to SW:XP\_A\_HUMAN P23025  
DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS ; mRNA sequence.  
B1792983  
ACCESSION B1792983.1 GI:15820708  
VERSION EST.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 598)  
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemisha, L., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blisstein, A.,  
Schiller, A., Theising, B., Rutter, E., Ronko, I., Bennett, J.,  
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Taagereishvili, R.,  
Williams, T., Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:

TITLE  
JOURNAL  
COMMENT

Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Juliana Brown  
(brown@fas.harvard.edu) This sequence now available from the IMAGE  
consortium, for clone orders contact: info@image.llnl.gov  
Seq primer: -40UP from Glibco  
High quality sequence stop: 451.  
Location/Qualifiers

# FEATURES

source

```
1..598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5670113"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_11b="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not I;
Site 2: Sal I; Starting library constructed using
Superscript Plasmid library kit (Life Technologies). cDNA
made by oligo-dt priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."
```

# ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 598;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCCATCTCATGTGTGATG 20  
|||||  
Db 354 GGTCCATCTCATGTGTGATG 373

RESULT 22  
BF446397 627 bp mRNA linear EST 01-DEC-2000  
LOCUS BF446397  
DEFINITION 7p36g05.x1 NCI CGAP P128 Homo sapiens cDNA clone IMAGE:3648176 3'  
similar to SW:XP\_A\_HUMAN P23025 DNA-REPAIR PROTEIN COMPLEMENTING  
XP-A CELLS ;contains MER37.t2 MER37 repetitive element ; mRNA  
sequence.  
BF446397  
ACCESSION BF446397.1 GI:11511535  
VERSION EST.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 627)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov  
Seq primer: -40UP from Glibco  
High quality sequence stop: 460.

TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
Location/Qualifiers  
1..627  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3648176"  
/sex="male"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1lb="NCI\_CGAP\_Pr28"  
/note="Organ: Prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneids 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaudo."

ORIGIN  
Query Match 100.0%; Score 20; DB 10; Length 627;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20  
|||||  
353 GGTCCTACTCATGTTGATG 372

RESULT 23  
BM853017/c 638 bp mRNA linear EST 06-MAR-2002  
LOCUS K-EST0134360 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-131-D04  
DEFINITION 5', mRNA sequence.  
ACCESSION BM853017  
VERSION BM853017.1 GI:19209416  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 638)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
CONTACT: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 131 row: D column: 04  
High quality sequence stop: 638.  
Location/Qualifiers  
1..638  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S22SNU16n1-131-D04"  
/sex="F"  
/tissue\_type="ascites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-16"  
/lab\_host="DH10B"  
/clone\_1lb="S22SNU16n1"  
/note="Organ: Stomach; Vector: pT7T3-Pac; Site\_1: EcoRI; Site\_2: NotI; The S22SNU16 library was contributed by the Soares laboratory and it was constructed as described by Bonaudo, M.F., Lennon, G. and Soares, M.B. (1996), Genome

Research 6(9): 791-806. RNA was prepared from harvested cells of SNU-16 culture. SNU-16 cell was obtained from Korean Cell Line Bank (KCLB). SNU-16 was established from ascitic fluids of Korean patients by Park J.G. et al. (1990), Cancer Res 50: 2773-2780."

ORIGIN  
Query Match 100.0%; Score 20; DB 12; Length 638;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20  
|||||  
69 GGTCCTACTCATGTTGATG 50

RESULT 24  
BI911749/c 639 bp mRNA linear EST 16-OCT-2001  
LOCUS BI911749  
DEFINITION BI911749 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5214347.5', mRNA sequence.  
ACCESSION BI911749  
VERSION BI911749.1 GI:16175618  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 639)  
NIH-MGC http://mgi.nci.nih.gov/.  
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)  
UNPUBLISHED (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgs@bbs-rcmail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
Plate: LNL11538 row: F column: 12  
High quality sequence start: 3  
High quality sequence stop: 632.  
Location/Qualifiers  
1..639  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5214347"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/clone\_1lb="NIH\_MGC\_118"  
/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 100.0%; Score 20; DB 12; Length 639;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTCCTACTCATGTTGATG 20  
|||||  
619 GGTCCTACTCATGTTGATG 600

RESULT 25  
 CB305774 672 bp mRNA linear EST 04-MAR-2003  
 LOCUS UI-CF-EN1-aeg-k-02-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone  
 DEFINITION UI-CF-EN1-aeg-k-02-0-UI 3', mRNA sequence.  
 ACCESSION CB305774  
 VERSION CB305774.1 GI:28846285  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 672)  
 REFERENCE Bonaldo, M.P., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 TITLE discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul.mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.reagen.com) or from Open Biosystems  
 (www.openbiosystems.com).  
 The following repetitive elements were found in this cDNA  
 sequence: 495-521, >AT rich#Low\_complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=yes.

FEATURES  
 source  
 1..672  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EN1-aeg-k-02-0-UI"  
 /tissue\_type="Primary Lung Cystic Fibrosis Epithelial  
 Cells"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_1ib="UI-CF-EN1"  
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-CF-EN1 is a normalized cDNA library containing the  
 following tissue(s): Primary Lung Cystic Fibrosis  
 Epithelial Cells. The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was ligated to an EcoR I adaptor, digested with Not  
 I, and cloned directionally into pT73-Pac vector. The  
 oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (GT)<sub>18</sub> tail. The  
 sequence tag for this library is CTGCTCAGCT.  
 TAG\_TISSUE=Human Lung Epithelial Cell lines untreated LPS  
 6hr to LPS 24h  
 TAG\_Lib=UI-CF-EN1  
 TAG\_SEQ=CTGCTCAGCT"

ORIGIN  
 Query Match 100.0%; Score 20; DB 14; Length 672;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGATG 20

Db 621 GGTCATCTCATGTGATG 640  
 |||||  
 RESULT 26  
 BM992090 678 bp mRNA linear EST 17-JUN-2002  
 LOCUS UI-H-DP1-auf-m-16-0-UI.s1 NCI CGAP\_DP1 Homo sapiens cDNA clone  
 DEFINITION IMAGE:5868855 3', mRNA sequence.  
 ACCESSION BM992090  
 VERSION BM992090.1 GI:19711479  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 678)  
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 TITLE Unpublished (1997)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 The following repetitive elements were found in this cDNA  
 sequence: 501-527, >AT rich#Low\_complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=yes.

FEATURES  
 source  
 1..678  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5868855"  
 /tissue\_type="Subchondral Bone"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_1ib="NCI CGAP DP1"  
 /note="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI CGAP\_DP1 is a normalized cDNA library containing the  
 following tissue(s): Subchondral Bone. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)<sub>18</sub> tail. The sequence tag for this library is  
 GTTAAGCTC.  
 TAG\_TISSUE=subchondral bone  
 TAG\_Lib=UI-H-DP1  
 TAG\_SEQ=GTTAAGCTC"

ORIGIN  
 Query Match 100.0%; Score 20; DB 12; Length 678;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCATCTCATGTGATG 20  
 |||||  
 Db 627 GGTCATCTCATGTGATG 646

RESULT 27  
 CA450156

LOCUS CA450156 705 bp mRNA linear EST 08-NOV-2002  
DEFINITION UI-CF-FNO-aff-k-13-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone  
ACCESSION CA450156  
VERSION CA450156.1 GI:24814576  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 705)  
AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PubMed 8895548  
COMMENT Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul.mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com) or from Open Biosystems  
(www.openbiosystems.com).  
The following repetitive elements were found in this CDNA  
sequence: 499-525, >AT rich#low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLY(A)=yes.  
FEATURES  
source  
Location/Qualifiers  
1..705  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-FNO-aff-k-13-0-UI"  
/tissue\_type="Human Lung Epithelial cells"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-CF-FNO is a subtracted cDNA library derived from two  
normalized Human Lung epithelial cell libraries (EN1 and  
DU1) The library was subtracted according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. For additional information, contact:  
bento-soares@uiowa.edu  
TAG\_TISSUE=Lung Epithelial Cells Tissue nos 359-368  
TAG\_LIB=UI-CF-FNO  
TAG\_SEQ=GGCTGAGGC"

ORIGIN  
Query Match 100.0%; Score 20; DB 14; Length 705;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCATCTCATGTTGATG 20  
|||||  
Db 625 GGTCATCTCATGTTGATG 644

RESULT 28 BQ772342 748 bp mRNA linear EST 26-JUL-2002  
LOCUS BQ772342  
DEFINITION UI-H-EZ1-bd1-1-20-0-UI.s1 NCI CGAP Ch2 Homo sapiens cDNA clone  
ACCESSION UI-H-EZ1-bd1-1-20-0-UI 3', mRNA sequence.  
BQ772342  
VERSION BQ772342.1 GI:21980818

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 748)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strassberg, Ph.D.  
Email: cga@db-remail.nih.gov  
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of  
Orthopedics  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@uiowa.edu  
The following repetitive elements were found in this CDNA  
sequence: 499-525, >AT rich#low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLY(A)=yes.  
FEATURES  
source  
Location/Qualifiers  
1..748  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-EZ1-bd1-1-20-0-UI"  
/tissue\_type="Chondrosarcoma Grade II"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia)  
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
NCI-CGAP\_Ch2 is a normalized cDNA library containing the  
following tissue(s): Chondrosarcoma Grade II. The library  
was constructed according to Ronaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pT73-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
(dfr)18 tail. The sequence tag for this library is  
TGATCAGCCT.  
TAG\_TISSUE=grade-2-chondrosarcoma  
TAG\_LIB=UI-H-EZ1  
TAG\_SEQ=ATCTAATATG"

ORIGIN  
Query Match 100.0%; Score 20; DB 13; Length 748;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGTCATCTCATGTTGATG 20  
|||||  
Db 625 GGTCATCTCATGTTGATG 644

RESULT 29 BG186679 790 bp mRNA linear EST 21-APR-2001  
LOCUS BG186679/c  
DEFINITION RSTS654 Athersys RACE Library Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG186679  
VERSION BG186679.1 GI:13708366  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 790)

**AUTHORS**  
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramchandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Maye, R., Smith, B., Veloso, N., Klika, A., Hess, J., Cochran, K., Lo, K., Ofensbacher, J., Panzig, J., and Ducar, M.  
Creation of genome-wide protein expression libraries using random activation of gene expression

**JOURNAL**  
Medline  
21227151  
Nat. Biotechnol. 19 (5), 440-445 (2001)

**TITLE**  
11329013

**COMMENT**  
Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave., Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@athersys.com  
High quality sequence stop: 353.  
Location/Qualifiers

**FEATURES**  
source  
1..790  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

**ORIGIN**  
Query Match 100.0%; Score 20; DB 12; Length 790;  
Best Local Similarity 100.0%; Pred. No. 96;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Db**  
1 GGTCCTACTCATGTTGATG 20  
255 GGTCCTACTCATGTTGATG 236

**RESULT 30**  
AY414079/c 809 bp DNA linear GSS 17-DEC-2003  
LOCUS  
DEFINITION Pan troglodytes XPA gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY414079  
VERSION AY414079.1 GI:39770041  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
1 (bases 1 to 809)  
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
Infering nonneutral evolution from human-chimp-mouse orthologous gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 809)  
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
FEATURES Location/Qualifiers

**source**  
1..809  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

**gene**  
1..809  
/gene="XPA"  
/locus\_tag="HCM5104"

**ORIGIN**  
Query Match 100.0%; Score 20; DB 29; Length 809;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Db**  
1 GGTCCTACTCATGTTGATG 20  
743 GGTCCTACTCATGTTGATG 724

**RESULT 31**  
AY414078/c 813 bp DNA linear GSS 17-DEC-2003  
LOCUS  
DEFINITION Homo sapiens XPA gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY414078  
VERSION AY414078.1 GI:39770040  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 813)  
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
Infering nonneutral evolution from human-chimp-mouse orthologous gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 813)  
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
FEATURES Location/Qualifiers  
source  
1..813  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

**gene**  
1..813  
/gene="XPA"  
/locus\_tag="HCM5104"

**ORIGIN**  
Query Match 100.0%; Score 20; DB 29; Length 813;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Db**  
1 GGTCCTACTCATGTTGATG 20  
743 GGTCCTACTCATGTTGATG 724

**RESULT 32**  
AL545429 873 bp mRNA linear EST 31-MAY-2003  
LOCUS  
DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS01027YAA03 5-PRIME, mRNA sequence.

ACCESSION AL545429  
VERSION AL545429.2 GI:31267264  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 873)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12877910.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4910.r For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI027AA020P1&cluster=4910.r. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CSODI027AA020P1.  
Location/Qualifiers  
1..873  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODI027YAO3"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 GGTCCATCTCATGTTGATG 20  
|||||  
809 GGTCCATCTCATGTTGATG 790

RESULT 33  
BM468458 1059 bp mRNA linear EST 05-FEB-2002  
LOCUS AGENCOURT 6432269 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5535843  
DEFINITION 5', mRNA sequence.  
ACCESSION BM468458  
VERSION BM468458.1 GI:18517500  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1059)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Place: LLNL12324 row: 3 column: 04

FEATURES  
source  
High quality sequence stop: 662.  
Location/Qualifiers  
1..1059  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5535843"  
/tissue\_type="lelomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NIH\_MGC\_71"  
/note="Orogen: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN  
Query Match 100.0%; Score 20; DB 12; Length 1059;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 GGTCCATCTCATGTTGATG 20  
|||||  
743 GGTCCATCTCATGTTGATG 724

RESULT 34  
AL573331 1114 bp mRNA linear EST 31-MAY-2003  
LOCUS AL573331 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CSODI042YC03 3-PRIME, mRNA sequence.  
ACCESSION AL573331  
VERSION AL573331.2 GI:31294686  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1114)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 16, 2001 this sequence version replaced gi:12932473.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4910.r For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI042AB02NP1&cluster=4910.r. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CSODI042AB02NP1.  
Location/Qualifiers  
1..1114  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODI042YC03"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 1114;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 GGTCCATCTCATGTTGATG 20  
|||||



Db 559 GGTCCTACTCATGTTGATG 578

RESULT 35  
AL548961/c 1129 bp mRNA linear EST 31-MAY-2003  
LOCUS AL548961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS0D1042YC03 5-PRIME, mRNA sequence.  
ACCESSION AL548961  
VERSION AL548961.2 GI:31270783  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 1129)  
Li W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:12884484.  
Contact: Genoscope - Centre National de Sequencage  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4910.r For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1042AB02QPl&cluster=4910.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1042AB02QPl.

FEATURES  
source  
location/Qualifiers  
1..1129  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1042YC03"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 1129;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
|||||  
Db 818 GGTCCTACTCATGTTGATG 799

RESULT 36  
B0879392/c 1174 bp mRNA linear EST 16-AUG-2002  
LOCUS B0879392 AGENCOURT 8118253 lupski\_dorsal\_root ganglion Homo sapiens cDNA  
DEFINITION clone IMAGE:6181232 5', mRNA sequence.  
ACCESSION B0879392  
VERSION B0879392.1 GI:22271400  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 1174)  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM13565 row: e column: 09  
High quality sequence start: 85  
High quality sequence stop: 369.  
location/Qualifiers  
1..1174  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6181232"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult, 36 yr"  
/lab\_host="DH10B"  
/clone\_lib="lupski\_dorsal\_root ganglion"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCGAGCGGCGCG-3' and 5'-GACTGTTCTGATGCGAGCGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 1174;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCCTACTCATGTTGATG 20  
|||||  
Db 242 GGTCCTACTCATGTTGATG 223

RESULT 37  
AL531628 1201 bp mRNA linear EST 23-MAY-2003  
LOCUS AL531628 Homo sapiens FETAL LIVER Homo sapiens cDNA clone  
DEFINITION CS0DM002YF23 3-PRIME, mRNA sequence.  
ACCESSION AL531628  
VERSION AL531628.2 GI:31069460  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 1201)  
Li W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 13, 2001 this sequence version replaced gi:12795121.  
Contact: Genoscope - Centre National de Sequencage  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4910.r For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DM002C12NP1&cluster=4910.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DM002C12NP1.  
location/Qualifiers  
1..1201

FEATURES  
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDM002f2p23"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_1ib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

## ORIGIN

```

Query Match          98.0%; Score 19.6; DB 9; Length 1201;
Best Local Similarity 95.0%; Pred. No. 1.8e+02;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GGTCCATCTCATGTTGATG 20
    |||||
Db 563 GGTCCATCTCATGTTGATG 582

```

```

RESULT 38          640 bp mRNA linear EST 08-SEP-1999
AI990924          ws24f07.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2498149 3'
LOCUS             similar to gb:Di4533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS
DEFINITION        (HUMAN); mRNA sequence.

```

```

ACCESSION         AI990924
VERSION           AI990924.1 GI:5837821
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens

```

```

REFERENCE          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE             1 (bases 1 to 640)
                  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
                  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                  Tumor Gene Index
JOURNAL           Unpublished (1997)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov
                  Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
                  R. Emmert-Buck, M.D., Ph.D.
                  cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
                  Bonaldo, Ph.D.

```

```

FEATURES
source            cDNA library Arrayed by: Greg Lennon, Ph.D.
                  DNA Sequencing by: Washington University Genome Sequencing Center
                  Clone distribution: NCI-CGAP clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  www.bio.llnl.gov/bbrp/image/image.html
                  Seq primer: -40UP from Gibco
                  High quality sequence stop: 459.
                  Location/Qualifiers

```

```

1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2498149"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_1ib="NCI CGAP GC6"
/notes="Vector: pRT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI-CGAP-GC4 was prepared, and
as circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

```

## ORIGIN

```

Query Match          95.0%; Score 19; DB 9; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2 GTCCATCTCATGTTGATG 20
    |||||
Db 605 GTCCATCTCATGTTGATG 623

```

## RESULT 39

```

LOCUS             BE763751          467 bp mRNA linear EST 19-SEP-2000
DEFINITION        RCS-NT0053-140600-022-D04 NT0053 Homo sapiens CDNA, mRNA sequence.
ACCESSION         BE763751
VERSION           BE763751.1 GI:10193675
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens

```

```

REFERENCE          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE             1 (bases 1 to 467)

```

```

JOURNAL           Shotgun sequencing of the human transcriptome with ORF expressed
MEDLINE           sequence tags
PUBMED            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
                  20202663
                  10737800
COMMENT           Contact: Simpson A.J.G.
                  Laboratory of Cancer Genetics
                  Ludwig Institute for Cancer Research
                  Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                  Brazil
                  Tel: +55-11-2704922
                  Fax: +55-11-2707001
                  Email: asimpson@ludwig.org.br

```

```

FEATURES
source            This sequence was derived from the FAPESP/LICR Human Cancer Genome
                  Project. This entry can be seen in the following URL
                  http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2-RCS-NT0053-140
                  600-022-D04et3=2000-06-14et4=1)
                  Seq primer: puc 18 forward
                  High quality sequence start: 3
                  High quality sequence stop: 464.
                  Location/Qualifiers

```

```

1..467
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1ib="NT0053"
/notes="Organ: nervous tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORG875 PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

## ORIGIN

```

Query Match          92.0%; Score 18.4; DB 10; Length 467;
Best Local Similarity 95.0%; Pred. No. 4.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 GGTCCATCTCATGTTGATG 20
    |||||
Db 416 GGTCCATCTCATGTTGATG 435

```

RESULT 40  
 B0321319/c 568 bp mRNA linear EST 17-MAY-2002  
 DEFINITION QV4-CT0491-080800-341-a11 CT0491 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION B0321319  
 VERSION B0321319.1 GI:20929303  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 568)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.U., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 JOURNAL 20202663  
 MEDLINE 10737800  
 PUBMED  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV4ct2=QV4-CT0491-  
 080800-341-a11&ct3=2000-08-08&ct4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 45.  
 Location/Qualifiers  
 1..568  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="CT0491"  
 /note="Organ: colon; Vector: puc18; Site 1: Sma1; Site 2:  
 Sma1; A mini-library was made by cloning products derived  
 from ORESTES PCR (U.S. Letters Patent application No.  
 196,716 - Ludwig Institute for Cancer Research) profiles  
 into the pUC 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."

ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 568;  
 Best Local Similarity 95.0%; Pred. No. 4.7e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTGTGATG 20  
 |||||  
 566 GGTCCATCTCATGTGTGATG 547

RESULT 41  
 B057361  
 LOCUS B057361 NIBB Mochii normalized Xenopus laevis cDNA, mRNA  
 DEFINITION laevis cDNA clone XL034n06 3', mRNA sequence.  
 ACCESSION B057361  
 VERSION B057361.1 GI:17470155  
 KEYWORDS EST.  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 625)  
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and  
 Kohara, Y.  
 Expressed genes in X. laevis embryo  
 Unpublished (2001)  
 JOURNAL Contact: Tadasu Shin-i  
 Center for Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@ggenes.nig.ac.jp  
 The information of this clone is available through the following  
 URL.  
 http://xenopus.nibb.ac.jp.  
 Location/Qualifiers  
 1..625  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="XL034n06"  
 /tissue\_type="whole embryo"  
 /dev\_stage="stage 15"  
 /clone\_lib="NIBB Mochii normalized Xenopus neurula  
 library"

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 625;  
 Best Local Similarity 95.0%; Pred. No. 4.9e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGTCCATCTCATGTGTGATG 20  
 |||||  
 49 GGTCCATCTCATGTGTGATG 68

RESULT 42  
 A1634902  
 LOCUS A1634902 711 bp mRNA linear EST 26-APR-1999  
 DEFINITION t274c02.x1 NCI CGAP Panel Homo sapiens cDNA clone IMAGE:2294306 3'  
 similar to gp:DI4533 DNA-REPAIR PROTEIN COMPLEMENTING XP-A CELLS  
 (HUMAN); mRNA sequence.  
 ACCESSION A1634902  
 VERSION A1634902.1 GI:4686232  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 711)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-rc@mail.nih.gov  
 Life Technologies catalog #: 11548-013  
 DNA sequencing by: Washington University Genome Sequencing Center  
 cDNA distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E.B. Consortium/ILNIB at:  
 www-bio.livl.gov/bbrp/image/image.html  
 Seq primer: -40UP from Glibco  
 High quality sequence stop: 409.  
 Location/Qualifiers  
 1..711  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2294306"  
 /tissue\_type="adenocarcinoma"  
 /lab\_host="DH10B"

FEATURES  
 source

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 711;  
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGTCCATCTCATGTTGATG 20  
 |||||  
 180 GGCCCATCTCATGTTGATG 199

/clone.lib="NCI CGAP Pan1"  
 /note="Organ: pancreas; Vector: PCMV-SPORT6; Site: 1; Salt:  
 Site: 2; Not: Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 711;  
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGTCCATCTCATGTTGATG 20  
 |||||  
 180 GGCCCATCTCATGTTGATG 199

RESULT 43  
 BGI87563 854 bp mRNA linear EST 21-APR-2001  
 LOCUS RST691 Atherys RAGE library Homo sapiens cDNA, mRNA sequence.  
 DEFINITION BGI87563  
 ACCESSION BGI87563.1 GI:13709378  
 VERSION EST.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.  
 1 (bases 1 to 854)  
 Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,  
 Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,  
 Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,  
 Maye, R., Smith, E., Veloso, N., Kliska, A., Hess, J., Cothren, K., Lo, K.,  
 Offenbacher, J., Danzig, J., and Ducar, M.  
 Creation of genome-wide protein expression libraries using random  
 activation of gene expression

TITLE  
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
 MEDLINE 21227151  
 PUBMED 11329013  
 COMMENT Contact: Scott J. Cain  
 Atherys, Inc.  
 3201 Carnegie Ave, Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596  
 Email: scain@atherys.com  
 High quality sequence stop: 482.  
 Location/Qualifiers

FEATURES  
 source  
 1..854  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="HT1080"  
 /clone.lib="Atherys RAGE library"  
 /note="See 'Creation of Genome-wide Protein Expression  
 Libraries using Random Activation of Gene Expression',  
 Nature Biotechnology, in press. Note that even though the  
 cell type indicated is HT1080, since a random activation  
 method was used, these sequence tags are not necessarily  
 expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 854;  
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGTCCATCTCATGTTGATG 20  
 |||||  
 602 GGTCCTACTCATGTTGATG 621

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 854;  
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGTCCATCTCATGTTGATG 20  
 |||||  
 602 GGTCCTACTCATGTTGATG 621

RESULT 44  
 AL571156 857 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL571156  
 DEFINITION AL571156 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

ACCESSION  
 AL571156  
 VERSION AL571156.2 GI:1292558  
 EST.  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.  
 1 (bases 1 to 857)  
 Li, M. B., Gruber, C., Jessup, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On Feb 16, 2001 this sequence version replaced gi:12928170.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 4910.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1027AA02NP1&cluster=4910.r. Contact :  
 Peng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1027AA02NP1.

FEATURES  
 source  
 1..857  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1027AA03"  
 /cissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone.lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-cligo (dt)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with NotI and cloned into the NotI and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 857;  
 Best Local Similarity 95.0%; Pred. No. 5.6e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGTCCATCTCATGTTGATG 20  
 |||||  
 588 GGTCATCTCATGTTGATG 607

RESULT 45  
 CC448351 1024 bp DNA linear GSS 22-MAY-2003  
 LOCUS ZMMBBOC329P17r ZMMBBOC Zea mays subsp. mays genomic clone  
 DEFINITION ZMMBBOC329P17 3', genomic survey sequence.  
 ACCESSION CC448351  
 VERSION CC448351.1 GI:30995914  
 KEYWORDS GSS.  
 SOURCE Zea mays subsp. mays (maize)  
 ORGANISM Zea mays subsp. mays

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoidae; Andropogoneae; Zea.  
 1 (bases 1 to 1024)  
 Bharti, A. K., Young, S., Kavchok, S., Keizer, G., Bronzino, A. C.,  
 Rouzard, K., Fu, G., Yu, Y., Wing, R., and Messing, J.  
 Sequencing of the maize genome at PCR (2003b)  
 Unpublished (2003)  
 CONTACT: Bharti, A. K.  
 Dr. Joachim Messing's lab  
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers  
 University  
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA  
 Tel: 732 445 3801  
 Fax: 732 445 5735

RESULT 45  
 CC448351 1024 bp DNA linear GSS 22-MAY-2003  
 LOCUS ZMMBBOC329P17r ZMMBBOC Zea mays subsp. mays genomic clone  
 DEFINITION ZMMBBOC329P17 3', genomic survey sequence.  
 ACCESSION CC448351  
 VERSION CC448351.1 GI:30995914  
 KEYWORDS GSS.  
 SOURCE Zea mays subsp. mays (maize)  
 ORGANISM Zea mays subsp. mays

Email: bharti@wakeman.rutgers.edu  
Seq primer: SP6  
Class: BAC ends  
High quality sequence start: 66.  
Location/Qualifiers

## FEATURES

source

1..1024  
/organism="Zea mays subsp. mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/sub\_species="mays"  
/db\_xref="taxon:4578"  
/clone="ZMMBCC0329P17"  
/lab\_host="E. coli DH10B"  
/clone\_lib="ZMMBAC"  
/note="Vector: pTARBAC1.3; Site\_1: BamHI; Site\_2: BamHI"

## ORIGIN

Query Match 92.0%; Score 18.4; DB 28; Length 1024;  
Best Local Similarity 95.0%; Pred. No. 6e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## Qy

1 GGTCATCTCATGTTGATG 20  
|||||

## Db

759 GGTCATCTCATGTTGATG 778

## RESULT 46

AM646323 420 bp mRNA linear EST 26-APR-2001  
cm63e10.w1 Blackshear/Soares normalized Xenopus egg library Xenopus  
laevis cDNA clone PBX0162E10 5', mRNA sequence.

## ACCESSION

AM646323

## VERSION

AM646323.1 GI:7403809

## KEYWORDS

EST.

## ORGANISM

Xenopus laevis (African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Xenopus.

1 (bases 1 to 420)

Blackshear,P.J., Lai,W.S., Thorn,D.M., Kennington,E.A., Staffa,N.G.  
Jr., Moore,D.T., Bouffard,G.G., Beckstrom-Sternberg,S.M.,  
Touchman,J.W., Donald,M.F. and Soares,M.B.

The NIHES Xenopus maternal EST project: interim analysis of the  
first 13,879 ESTs from unfertilized eggs

Gene 267 (1), 71-87 (2001)

## TITLE

Contact: Perry J. Blackshear

## JOURNAL

Office of Clinical Research and Laboratory of Signal Transduction  
National Institute of Environmental Health Sciences

## MEDLINE

A2-05 NIHES, 101 Alexander Drive, Research Triangle Park, NC 27709,  
USA

## PUBMED

Tel: 919 541-4899  
Fax: 919 541-4571

## COMMENT

Email: black009@niehs.nih.gov  
Clone is available through Research Genetics, Inc., 2130 Memorial  
Parkway, Huntsville, AL 35901

## FEATURES

phone 800-533-4363 ext.cdna, fax 256-536-9016 att.cdna, email  
cdna@resgen.com

## source

DNA Sequencing and analyses performed by National Institutes of  
Health Intramural Sequencing Center (NISC).

## PCR Primers

FORWARD: TGTAAACGACGGCCAGT  
BACKWARD: CAGGAACACGCTATGACC

## Plate: 0162

row: E column: 10

## Seq primer: T7 primer.

Location/Qualifiers

## 1..420

/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="PBX0162E10"

## ORIGIN

Query Match 87.0%; Score 17.4; DB 10; Length 420;  
Best Local Similarity 94.7%; Pred. No. 1.2e+03;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## Qy

2 GGTCATCTCATGTTGATG 20  
|||||

## Db

202 GGTCATCTCATGTTGATG 220

## RESULT 47

BJ083380 492 bp mRNA linear EST 29-SEP-2003  
BJ083380 NIBB Mochii normalized Xenopus tailbud library Xenopus  
laevis cDNA clone XL085a05 3', mRNA sequence.

## ACCESSION

BJ083380

## VERSION

BJ083380.1 GI:17578422

## KEYWORDS

EST.

## ORGANISM

Xenopus laevis (African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Xenopus.

1 (bases 1 to 492)

Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and  
Kohara,Y.

Expressed genes in X. laevis embryo  
Unpublished (2001)

## TITLE

Contact: Tadasi Shin-i

## JOURNAL

Center For Genetic Resource Information  
National Institute of Genetics

## MEDLINE

1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855

## PUBMED

Email: tshini@genes.nig.ac.jp  
The information of this clone is available through the following  
URL.

## COMMENT

http://xenopus.nibb.ac.jp.  
Location/Qualifiers

## 1..492

/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="XL085a05"  
/tissue\_type="whole embryo"  
/dev\_stage="stage 25"  
/clone\_lib="NIBB Mochii normalized Xenopus tailbud  
library"

QY 1 GGTCATCTCATGTTGATG 20  
 |||||||||  
 Db 429 GGTCATCTCATGTTGATG 448

RESULT 48  
 AL637062/c 587 bp mRNA linear EST 07-NOV-2003  
 LOCUS AL637062 XGC-neurula Silurana tropicalis cDNA clone TNeu003d20 5',  
 DEFINITION mRNA sequence.  
 ACCESSION AL637062 GI:38215704  
 VERSION AL637062.2 GI:38215704  
 KEYWORDS EST.  
 SOURCE Silurana tropicalis (western clawed frog)  
 ORGANISM Silurana tropicalis  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 Xenopodinae; Silurana.  
 1 (bases 1 to 587)  
 REFERENCE Croining,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
 AUTHORS Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
 JOURNAL Unpublished (2003)  
 COMMENT On Nov 7, 2001 this sequence version replaced gi:16789041.  
 Contact: Huckle E  
 Sanger Institute  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Aaron M. Zorn.  
 cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.  
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the  
 5' end and NotI at the 3' end.  
 Vector: pCS107; Site\_1: EcoRI; Site\_2: NotI  
 Host: Escherichia coli DH10B  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS\_SEQUENCE\_ID: TNeu003d20.plksp6  
 Sequencing primer: Sp6.  
 FEATURES  
 source location/Qualifiers  
 1..587  
 /organism="Silurana tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="TNeu003d20"  
 /dev\_stage="neurula"  
 /lab\_host="Escherichia coli DH10B"  
 /clone\_lib="XGC-neurula"  
 /note="Vector: pCS107; Site\_1: EcoRI; Site\_2: NotI; cDNA  
 was oligo dt primed from 5ug of poly A+ RNA from neurula.  
 EcoRI-NotI cut cDNA was then ligated into pCS107 with  
 EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 587;  
 Best Local Similarity 94.7%; Pred. No. 1.4e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20  
 |||||||||  
 Db 507 GTCCATCTCATGTTGATG 489

RESULT 49  
 CD434346 828 bp mRNA linear EST 03-JUN-2003  
 DEFINITION EL01N0333A02.b Endosperm\_3 Zea mays cDNA, mRNA sequence.  
 ACCESSION CD434346  
 VERSION CD434346.1 GI:31349989  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

REFERENCE  
 AUTHORS lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and  
 Messing,J.  
 TITLE Sequencing of the maize endosperm ESTs  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: lai, jinsheng  
 Dr. Joachim Messing's lab  
 Waksman Institute, Rutgers University  
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA  
 Tel: 732-445-3801  
 Fax: 732-445-5735  
 Email: jlai@waksman.rutgers.edu  
 Seq primer: T3.  
 FEATURES  
 source location/Qualifiers  
 1..828  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="W22"  
 /db\_xref="taxon:4577"  
 /tissue\_type="Endosperm of 7-23DAP"  
 /clone\_lib="Endosperm 3"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI"

ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 828;  
 Best Local Similarity 94.7%; Pred. No. 1.6e+03;  
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GTCCATCTCATGTTGATG 20  
 |||||||||  
 Db 296 GTCCATCTCATGTTGATG 314

RESULT 50  
 B1261004 808 bp mRNA linear EST 17-JUN-2001  
 LOCUS B1261004  
 DEFINITION 602972177F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:511539 5',  
 mRNA sequence.  
 ACCESSION B1261004 GI:14819850  
 VERSION B1261004.1 GI:14819850  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 808)  
 REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.  
 AUTHORS NIH-MGC  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgs@bbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLAN at:  
 http://image.llnl.gov  
 Plate: L1M11270 row: j column: 20  
 High quality sequence start: 63  
 High quality sequence stop: 173.  
 FEATURES  
 source location/Qualifiers  
 1..808  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:511539"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_12"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site\_1: NotI;

Site 2: Sail; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.4 kb. Library prepared by Life  
Technologies."

## ORIGIN

Query Match 85.0%; Score 17; DB 12; Length 808;  
Best Local Similarity 100.0%; Pred. No. 2.5e+03;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 GTCCATACCTCAGTTGA 18  
Db 48 GTCCATACCTCAGTTGA 64

Search completed: August 17, 2004, 15:45:29  
Job time : 2606 secs

**THIS PAGE BLANK (USPTO)**